

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 09:19:57 ; Search time 23 Seconds
(without alignments)
1087.280 Million cell updates/sec

Title: US-10-724-972A-6352
Perfect score: 1690
Sequence: 1 GVESVRLGKILSVIGLLFVL.....EEMAKELVELSKDCKDK 335

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep: *
2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep: *
3: /cgn2_6/ptodata/1/1aa/6A-COMB.pep: *
4: /cgn2_6/ptodata/1/1aa/6B-COMB.pep: *
5: /cgn2_6/ptodata/1/1aa/PTUS-COMB.pep: *
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep: *

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	405	24.0	330	4	US-09-830-217-16
2	405	24.0	330	4	US-08-956-171B-5193
3	405	24.0	330	4	US-08-781-986A-5193
4	405	24.0	330	4	US-10-278-946-16
5	307.5	18.2	352	4	US-09-107-532A-4516
6	296.5	17.5	351	3	US-09-134-001C-5492
7	296	17.5	351	4	US-09-543-681A-6862
8	292	17.3	347	4	US-09-710-279-462
9	292	17.3	347	4	US-09-710-279-1268
10	283	16.7	317	4	US-09-071-035-38
11	283	16.7	325	4	US-09-134-000C-4342
12	278.5	16.5	325	4	US-09-543-681A-5565
13	268	15.9	296	4	US-09-071-035-40
14	266	15.7	306	4	US-09-602-787A-572
15	265	15.7	60	3	US-09-134-001C-4860
16	254	15.0	344	4	US-09-583-110-3723
17	250	14.8	321	4	US-09-583-110-2973
18	250	14.8	325	4	US-09-107-433-4809
19	246.5	14.6	332	4	US-09-602-787A-598
20	241	14.3	324	4	US-09-328-352-4444
21	240	14.2	312	4	US-09-035-70
22	236	14.0	298	3	US-08-961-083-24
23	236	14.0	298	4	US-09-536-784-24
24	236	14.0	301	4	US-09-134-000C-6014
25	228	13.5	289	4	US-09-071-035-72
26	218	12.9	321	4	US-09-302-626B-12
27	217	12.8	321	4	US-09-302-626B-10

ALIGNMENTS

RESULT 1

US-09-830-217-16
; Sequence 16, Application US/09830217
; Patent No. 6521441
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Staphylococcus aureus Genes and Polypeptides
; FILE REFERENCE: PB461PCT
; CURRENT APPLICATION NUMBER: US/09/830,217
; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: PCT/US99/06199
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: 60/080,296
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/084,674
; PRIOR FILING DATE: 1998-05-07
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 16
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-830-217-16

Query Match	24.0%	Score 405;	DB 4;	Length 330;
Best Local Similarity	32.8%	Pred. No. 1.9e-25;		
Mismatches	113;	Conservative	65;	Mismatches 132;
Indels	34;	Gaps	12;	
QY	8	LKILSVIGLLFVLIATAACGNSSSSKSSKGGVEIKHEGTTKVPKPKRVVTVLEYS	67	
Db	5	IKML-VVTLALFLV-LACGSGNSNQSDNDKDKTTSIKHANGTTEIKGPKRVVTVLQV	62	
QY	68	FVDALVALDVPVGIADNDKKNRIIKLRDKIGKTVTSVTRKQPNLEISKLKPLIIAD	127	
Db	63	ATDVAVSLGVKPVGAVESWTQPKPEYIKNDLKDTPKIVGQEPAPNLEISKLKPLIIAS	122	
QY	128	NNRHGIVKDKLNIAPTIELSKFDGDYNNIDAFKTSIKALGKEEGKKRLEEHDKKIE	187	
Db	123	KVRNEKVDQLSKIAPTV---STDVTFKFK-DTTLKMGKALGKEAEADLLKKYDDKVA	178	
QY	188	YKKEITMDKNQKVLPAVAAGSLLAHPSN-----SYVGQFLSOLGPKKEALSDVDVKGL	240	
Db	179	FOK-----DAKAKYKDAWPLKASVNNFRADHTRIYAGGATAGETLNDLGFKR-----NKDL	228	
QY	241	SKYLKG--PYLQMT-ETLSQVNPBRMFTMTNKASNEPSL-----KELEKDPVVMKCLNA	292	
Db	229	QKQVNGKDIILTSTKESIPLMNADHIFVKSDPNKDAALVKKTESEWTSSKEWQNLDA	288	
QY	293	VKNQRV-DILDRDLNWRSGRLISSEMAKELVELSKDCKDK 335		

Db 289 VKNQVSDDLDEITWNLAGGYKSSLSKLDDLYE--KLNIEKQSK 330

RESULT 2

US-08-956-171E-5193
; Sequence 5193, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hymen
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 5193:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 330 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5193:
US-08-956-171E-5193

Query Match 24.0%; Score 405; DB 4; Length 330;
Best Local Similarity 32.8%; Pred. No. 1.9e-25;
Matches 113; Conservative 65; Mismatches 132; Indels 34; Gaps 12;

Qy 8 LKILSVIGLLFVLVLTATAACGNSSSSSKSGDVEIKHEEGTTKVPKPKRVVVLYQS 67
Db 5 IKML-VVTLAFLV-LAGCSGNSKQSSDNKDKEITTSIKHAMGTTTEIKGPKRVVTLVQ 62

Qy 68 FVDALVALDVPVGIADNKNRIIKPLRDKIGKYSVGTGRKQPNLEISKLKPDLIAD 127
Db 63 ATDVAVSLGVKPVGAVESWTQKPFYIKNDLKDTKIVGQBPAPNLEISKLKPDLIAS 122

Qy 128 NNRHKGIVKDLNKTAPTIETLSKSGDYNENIDAFKTSKALGKEEGKRLKEHDKKTEE 187
Db 123 KVRNEKYDQLSKAPTIV---STDVTFKFK-DTTKLMGKALGKEKEAEDLLKKYDDKVA 178

Qy 188 YKKEITMDKNQKVLPAVAASKSGLLAHPSN-----SYVGQFLSQLGFEALSDDVTKGL 240
Db 179 FQK----DAKAKYKDAWPLKASVNVNFRADHTRIYAGGVAGEILNDLGFKR-----NKDL 228

Qy 241 SKYLGK--PYLOMNT-ETLSQVNPFRMFMINKASSNEPSL-----KELEKDPVWKKLA 292
Db 229 QKQVDNGKDIQITSKESIPLMNADHIFVVKSDPNKADALVVKTEBWTSSKEWKNLDA 288

Qy 293 VKNQRV-DILDRDLWARSRLISSEEMAKELVELSKDKSKDNK 335
Db 289 VKNQVSDDLDEITWNLAGGYKSSLSKLDDLYE--KLNIEKQSK 330

RESULT 3

US-08-781-986A-5193
; Sequence 5193, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 5193:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 330 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-781-986A-5193

Query Match 24.0%; Score 405; DB 4; Length 330;
Best Local Similarity 32.8%; Pred. No. 1.9e-25;
Matches 113; Conservative 65; Mismatches 132; Indels 34; Gaps 12;

Qy 8 LKILSVIGLLFVLVLTATAACGNSSSSSKSGDVEIKHEEGTTKVPKPKRVVVLYQS 67
Db 5 IKML-VVTLAFLV-LAGCSGNSKQSSDNKDKEITTSIKHAMGTTTEIKGPKRVVTLVQ 62

Qy 68 FVDALVALDVPVGIADNKNRIIKPLRDKIGKYSVGTGRKQPNLEISKLKPDLIAD 127
Db 63 ATDVAVSLGVKPVGAVESWTQKPFYIKNDLKDTKIVGQBPAPNLEISKLKPDLIAS 122

Qy 128 NNRHKGIVKDLNKTAPTIETLSKSGDYNENIDAFKTSKALGKEEGKRLKEHDKKTEE 187
Db 123 KVRNEKYDQLSKAPTIV---STDVTFKFK-DTTKLMGKALGKEKEAEDLLKKYDDKVA 178

Qy 188 YKKEITMDKNQKVLPAVAASKSGLLAHPSN-----SYVGQFLSQLGFEALSDDVTKGL 240
Db 179 FQK----DAKAKYKDAWPLKASVNVNFRADHTRIYAGGVAGEILNDLGFKR-----NKDL 228


```

RESULT 6
US-09-134-001C-5492
; Sequence 5492, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5492
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5492

Query Match 17.5%; Score 296.5; DB 3; Length 351;
Best Local Similarity 28.4%; Pred. No. 1.8e-16;
Matches 104; Conservative 70; Mismatches 131; Indels 61; Gaps 16;

Qy 1 GVESVRGLKILSVIGLLFVLIIATAACGNSSSSSKES-SKDGVEIKHE----- 48
Db 1 GVAHMKTVLPFLLSLVILVTACSNSSNNSTSKKNSDSKETVTKNSFEASCKENNGS 60
Qy 49 -----EGTTKVPKHPKRVVLEYSFVDALVALDVKPVGIADDNK-----KNRIIKPLRD 97
Db 61 DKXKISNTVEVPKPKNAVLDYDALDVL-----KELGVADKVKGLPKGENNQSLPKFLD 115
Qy 98 --KIGKTSYVGTQKPNLEESIKLKPDI-IADNNRHKGIYKDLNKIAPTIELASFPDGY 154
Db 116 EFKDDKYINTGLNKLEVNFPDKVASAKPDVIFISGRATQKNLDEPKKAPKAKV-VYVGTG 174
Qy 155 NENI--DAFK---TISKALGKEESGKKRLEHDKKIBEYKKEITMDKKNQKVLPAVAAKSG 209
Db 175 DDNLIKDMKNTENGLKIYKEDKAKKINKDLDRKISDMKDK-KPDKFKVMYLLVNEGE 233
Qy 210 LLAH-PNSVYVQGLSOLGPKFALSDDVTGKLSKYLKGPYLQ-MNTETLSQNP----- 261
Db 234 LSTFGPGRFGGLVFDTLGPKFA-----DKKVSXSPHQINNEYINKQNPVDILAM 285
Qy 262 ERMFIMTNKASSNPSPSLKELEKDPWKWKLNAVKNQRVDILDRDLWARSGLISSEEMAKE 321
Db 286 DRGSVVGKATTN-----QVLKNKVIKNVKAVKSNNHIYELDPKLMWYFSSG--SSTTTIKQ 338

Qy 322 LVELSK 327
Db 339 IDELNE 344

RESULT 7
US-09-543-681A-6862
; Sequence 6862, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6862
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Proteus mirabilis

```


QY 221 QFLSOLGPKFKEALSDDVTGKLSKYLKGPYLQ-MNTETLSQVNP-----ERMFIMTNKASS 273
DB 242 LVFDTLGFKPA-----DKKVSQSPHGQINNEYINQKQPDVILAMDGRSVVGGKATT 293
QY 274 NEPSLKELEKDPVWKLANVQNRVDIILDRDLWARSGLISSEEMAKELVELSK 327
DB 294 N-----QVLKKNKVIKNAVKSNHIYELDPKLMYFSSG--SSTTTIKQIDELNE 340

RESULT 9

US-09-710-279-1268
; Sequence 1268, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1268
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-1268

Query Match 17.3%; Score 292; DB 4; Length 347;
Best Local Similarity 28.5%; Pred. No. 4.1e-16;
Matches 101; Conservative 68; Mismatches 121; Indels 64; Gaps 16;

QY 16 LLFVLVLTAAACGNSSSSSKE-----SSKDGVEIKHE-----EGTKVP 55
DB 9 LLSLVLTACSNSSNNSTSKKNSDSKETVTIKNSFEASKENNGSDKKISNTEVP 68
QY 56 KHPKRVVLEYSFVDALVALDVKPGIADDNK-----KNRIKPLRD--KIGKTSYVGT 107
DB 69 KNPKNVLDYGLDVL-----KELGVAUKVGLPKGNNQSLPKFLDFDKDKYINTGN 123
QY 108 RKOPNLETSKLPDLI--TADNNRHGIYKDLNKAFTIELSKSPDGDYNNENI--DAFK-- 162
DB 124 LKEVNFDKVASAPDVFISGRANQNLDEFKKAAPKAKV-VYVGTSDNLIKDKMKNT 182
QY 163 -TISKALGKEEGKKELEHDKKIEYKKEITMDKNQKVLPAVAASKGLLAH-PSNSYVG 220
DB 183 ENLGKIYDKEDKAKKINKOLDRIKSDMKDK-TKDFNKVYLLVNEGELSTFGPGRFG 241
QY 221 QFLSOLGPKFKEALSDDVTGKLSKYLKGPYLQ-MNTETLSQVNP-----ERMFIMTNKASS 273
DB 242 LVFDTLGFKPA-----DKKVSQSPHGQINNEYINQKQPDVILAMDGRSVVGGKATT 293
QY 274 NEPSLKELEKDPVWKLANVQNRVDIILDRDLWARSGLISSEEMAKELVELSK 327
DB 294 N-----QVLKKNKVIKNAVKSNHIYELDPKLMYFSSG--SSTTTIKQIDELNE 340

RESULT 10

US-09-071-035-38
; Sequence 38, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue

; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 317 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-071-035-38

Query Match 16.7%; Score 283; DB 4; Length 317;

Best Local Similarity 30.4%; Pred. No. 2e-15;
Matches 101; Conservative 59; Mismatches 130; Indels 42; Gaps 13;
QY 11 LSVIGLFLVLIATAACGN--SSSNSKSSKSGKGVKHEEGTTKVPKHPKRVVVLAYS 67
DB 11 VSMGMLML-----SACQTKKTADSATTETAKTEVTVDKTNQGLTPKPKVVFVDFG 66
QY 68 FVDALVALDV--KPGIADDNKNRIKPLRDKIGKYTSVTRKOPNLEESKLPDLII 125
DB 67 SLDTMDALGVDRVVGAPTKN-----PAYLKVKQKVESAGIKPDLKINQLKAPDLII 121
QY 126 ADNRRHKGIYKDLNKAFTIEL-----KSFQGDYNNENIDAFKTIKALGKEEGKKEE 181
DB 122 I-SGQQDYQQLKAIAPTIYLVADKNPWASTKQNIETLGTI---FDKEEVAKKEKITGL 177
QY 182 DKKIEYKKEITMDKNQKVLPAVAASKGLLAHPSNSYVGGFLSOLGPKFKEALSDD-----VT 237
DB 178 EKEIADVKKQABASAN-NALVVLVNEGQLSAYGKSRFGLIHDTEGFKAA--DDKIEAST 234
QY 238 KGLSKYLKGPYLMNTETLSQVNPMEFIM-TNKASSNEPSLKELEKDPVWKLANVKNQ 296
DB 235 HGQS-----VSIEYVLEKNPGILFVVDRTKAIGDDSKDNVAANELIQTDAKND 285
QY 297 RVDIILDRDLWARSGLISSEEMAKELVELSK 328
DB 286 KVIMLPQDVWYLSGGGLESMLH---MIEDVK 314

RESULT 11

US-09-134-000C-4342
; Sequence 4342, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778

```
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4342
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4342

Query Match      16.7%; Score 283; DB 4; Length 327;
Best Local Similarity 30.4%; Pred. No. 2.1e-15;
Matches 101; Conservative 59; Mismatches 130; Indels 42; Gaps 13;

Qy 11 LSVIGLLFVLIIATAACGNN-----SSNSSKESKSGVEIKHEEGTTKVPKPKRVVVLEYS 67
Db 17 VSMGLML-----SACQTNKKTADSATTETAKTEVTVDKTNQGLTVPKPKRVVFDNG 72
Qy 68 FVDALVALDV--KPGVIADNKKRIIKPLRDKIGKTSVGRKOPNLEEISKLKPLII 125
Db 73 SLDTMDALGVDGVVVGAPTKN-----IPAYLKVKQVESAGGIKEPDLEKINQLKPDII 127
Qy 126 ADNRHKGIVKDLNKAFTIEL-----KSFQDGDYNNENIDAFKTSKALGKEEGKKRLEE 181
Db 128 I-SGRQQDYQQLKAIATPIYLAVDKAKNPWASTKQNIETLGTI---FDKEEVAKEKITGL 183
Qy 182 DKIEEYKEITMDKNQKVLPAVAAGSGLLAHPSNVYQGLSQGLQFKEALSDD-----VT 237
Db 184 EKEIADVKKQAEASAN-NALVVLVNEGQLSAYGKSGREFLIHDTFGFKAA--DDKIEAST 240
Qy 238 KGLSKYLKGPYLQWNTETLSQVNPFRMFI--TNKASSNEPSLKELEKDPVWKKLNAVKNQ 296
Db 241 HQQS-----VSYEVLKNGPILFVVDRTKAIGGDDSKDNVAANELIKTDAGKND 291
Qy 297 RVDILDRDLWARSRLISSEEMAKELVELSKK 328
Db 292 KVINLQPDVWYLSGGGLSEMLH---MIEDVKK 320

RESULT 12
US-09-543-681A-5565
; Sequence 5565, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETTON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5565
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5565

Query Match      16.5%; Score 278.5; DB 4; Length 325;
Best Local Similarity 28.0%; Pred. No. 4.9e-15;
Matches 94; Conservative 65; Mismatches 132; Indels 45; Gaps 11;

Qy 9 KILSVIGLLFVLIIATAACGNNSS-----SNSKESKSGVEIKHEEGTTKVPKPKRVVVL 64
Db 12 KILAVTASLLSVSFSALSNTSVYQPNQIVEQQNDKVKVILKHLIGETAVSKNPKSVLVF 71
Qy 65 EYSFVDALVALDVK-----PVGADNKKRIIKPLRDKIGKTSVGRKOPNLEEISK 118
Db 72 DFLGVDLSVLGLADKVVGLPLGNAPYIKRIANDV-----ANVGKMAPFERLAE 124
Qy 119 LKPDIIADNRHKGIVKDLNKAFTIELKPSDGDYNNENIDA-FKTSKALGKEEGKKR 177
Db 125 LKPDIIII-TGRQGSYEKLTNIAPTINLTGNSANYLSDSVEANIQLIGELFDKQQAQTQ 183
```

```
Qy 178 LEEHDKKIEBKKEITMDKNQKVLPAVAAGSGLLAHPSNVYQGLSQGLQFKEALSDDVT 237
Db 184 LATLNTVIEQAQKK-AAGSDKKVLLVLLHNAAGNLN---PNN-----QSVIYDVV 227
Qy 238 KGLSKYL-----KGPYLQWNTETLSQVNPFRMFI--TNKASSNEPSLKE--LEKDPVWK 288
Db 228 KAKKAELPPVAEEKGRVVTSEMIKAVPDVFIIDRSEATGAGKLEKNTFENDEV-K 286
Qy 289 KLNKVNQRVDILDRDLWARSRLISSEEMAKELVE 324
Db 287 TTSAYKGGIVYLQSLDWLWYLSGGGLSLTQQIEAVQ 322

RESULT 13
US-09-071-035-40
; Sequence 40, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brooks
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 296 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-071-035-40

Query Match      15.9%; Score 268; DB 4; Length 296;
Best Local Similarity 30.3%; Pred. No. 3.2e-14;
Matches 96; Conservative 54; Mismatches 129; Indels 38; Gaps 12;

Qy 26 CGNN-----SSNSSKESKSGVEIKHEEGTTKVPKPKRVVVLYSFVDALVALDV--KPV 80
Db 1 CQTNKKTADSATTETAKTEVTVDKTNQGLTVPKPKRVVFDNGSLDTMDALGVDGVV 60
Qy 81 GIADNKKRIIKPLRDKIGKTSVGRKOPNLEEISKLKPDIIADNRHKGIVKOLNK 140
Db 61 GAPTKN-----IPAYLKVKQVESAGGIKEPDLEKINQLKPDIIII-SGRQQDYQQLKA 114
Qy 141 IAPTIEL-----KSPDGDYNNENIDAFKTSKALGKEEGKKRLEEHDKKIEBKKEITMDK 196
Db 115 IAPTIIYLAVDKAKNPWASTKQNIETLGTI---FDKEEVAKEKITGLEKEIADVKQAEASA 171
```


; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4860
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4860

Query Match 15.7%; Score 265; DB 3; Length 60;
Best Local Similarity 92.9%; Pred. No. 6.3e-15;
Matches 52; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 280 ELEDKPVWKKLVNAVKQNRVDILDRDLWARSGLISSEEMAKELVELSKDKSKDKK 335
Db 5 QIRKRPVWKKLVNAVKQNRVDILDRDLWARSGLISSEEMAKELVELSKDKSKDKK 60

RESULT 16

US-09-583-110-3723
; Sequence 3723, Application US/09583110
; Patent No. 669703
; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: Pneumoniae for Diagnostics and Therapeutics
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322

; SEQ ID NO 3723
; LENGTH: 344
; TYPE: PRT

; ORGANISM: Streptococcus pneumoniae
US-09-583-110-3723

Query Match 15.0%; Score 254; DB 4; Length 344;
Best Local Similarity 28.9%; Pred. No. 5.5e-13;
Matches 93; Conservative 58; Mismatches 115; Indels 56; Gaps 15;

Qy 13 VIGLLFVLIAACGNNSSSSSSKESKDGVEIKHEEGTTKVPKHPKRVVLEYSFV 69
Db 10 LIATLACIVFSACSSSVKNEENTSKEHAPDKIVLDHAFGQTILDKKPERVATIANGNH 69
Qy 70 DALVALDVKPGVIADDN---KKNRIKP-----LRDKIGKVTSGTRKQPNLEISKLKP 121
Db 70 DVALALGIVPGFSKANYGVSAADKGVLPWTEKIKELNGKANLPDLDGLFAEALSNSKP 129
Qy 122 DLIADNRRHGI-----YKLNKIAPTIELKSPDGYNNENI--DAFTISKALGKEEG- 174
Db 130 DVILAG---YSGITKEDYDTLSKIAPVAAVAKSPW---QTLWRDMIKIDSKALGMEKEGD 183
Qy 175 -----KKRLEHDKTIEVKETIMDKNQKVLPAV---AAKSGLLAHPNSYVGQ 221
Db 184 ELINKTEARISKLEKH-----PEITGKI---KGKKVLPTMINAADTSKFWIYTSKDPNPN 236
Qy 222 FLSQLG--FKBALSDDVTGKLSKYLKPGYLOMNTETLSQVNPFRMFMITNKASSNEPSLK 279
Db 237 YLTDLGLVFPSPSLKEFSE--DSFAK-----EISAEANKINDADVII---TYGDDKTLE 286
Qy 280 ELEDKPVWKKLVNAVKQNRVDIL 301
Db 287 ALQKDPLLGKINAINGAVAVI 308

RESULT 17

US-09-583-110-2973
; Sequence 2973, Application US/09583110
; Patent No. 669703
; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: Pneumoniae for Diagnostics and Therapeutics
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322

; SEQ ID NO 2973
; LENGTH: 321
; TYPE: PRT

; ORGANISM: Streptococcus pneumoniae
US-09-583-110-2973

Query Match 14.8%; Score 250; DB 4; Length 321;
Best Local Similarity 27.2%; Pred. No. 1.1e-12;
Matches 92; Conservative 65; Mismatches 119; Indels 62; Gaps 14;

Qy 20 LIATAACGNNSSSSSSKESKDGVE--IKHEEGTTKVPKHPKRVVLEYSFVDALVALDV 77
Db 17 LLLLGACSTNSSTISQTSSTSSAPTEVTKSLDEVKLSKYPEKIVTDLGAADTIRALGP 76
Qy 78 KPGVIADNKNRIKP-----LRDKIGKVTSGTRKQPNLEISKLKPDLIADNRR 130
Db 77 -----EKNIVGMPTKTVTYKDLVGTKNVGFMPKEPDLEAIALEPDLIASPR 127
Qy 131 HKGIYKDLNKIAPTIELKSPDGYNNENIDA-FKTIISKALGK--BEEGKKRLEEHDKKEE 187
Db 128 QKFVDK-FKEIAPTVLFOAKDDYWTSTKANIESLASAFGTGTQKAKEELAKLDKSIQE 186
Qy 188 Y-----KKEITMDKNQKVLPAVAAKSGLLAHPNSYVGQFLSQ-LGPKAELSDDDVTK 238
Db 187 VATKNESDDKALAILNNEGMAAFGAKSRF-----SFLYQTLKFKPT---DTTF 233
Qy 239 GLSKYLKPGYLOMNTETLSQVNPFRMFMITNKA-----SSNEPSLKE--LEKDPVWKKL 290
Db 234 EDSRHGQ---EVSFESVKIINPDILFVINRTLAIGDSSNNNGVLENALIAETP----- 284
Qy 291 NAVGNQRVDILDRDLWARSGLISSEEMAKELVELSKK 328
Db 285 -AAKNGKIIQLTPDLWYLSGGGLESTKLMIEDIQALK 321

RESULT 18

US-09-107-433-4809
; Sequence 4809, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS

; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: <Unknown>

FOR DIAGNO

```

; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION NUMBER: US/09/107,433
; APPLICATION NUMBER: 60/085131
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION NUMBER: 60/085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneka
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 4809:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 325 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8) LOCATION 1...325
; SEQUENCE DESCRIPTION: SEQ ID NO: 4809:
US-09-107-433-4809
Query Match 14.8%; Score 250; DB 4; Length 325;
Best Local Similarity 27.2%; Pred. No. 1.1e-12;
Matches 92; Conservative 65; Mismatches 119; Indels 62; Gaps 14;
QY 20 LIATAACGNSSSSSKESKDGVE--IKHERGTTKVPKHPKRVVVVLEYSFVDALVALDV 77
Db 21 LLLGACSTNSSTQTSSTSSAPTEVTKSSLDVKLSKVPKIVTFDLGAADTIRALGF 80
QY 78 KPVGIADDNKNRIKTP-----LRDKIGKTVTSVTRKQPNLEBISKLPDLIIADNRR 130
Db 81 -----EKNIIVGPTKTVPTYLKDLVGTGVKNVGFMEKPEDLAALAEPLIIASPT 131
QY 131 HGIIYKDLNKIAPTELKSPFDGYNENIDA-PKTIKALGK--EEGKKRLEEHDKLEE 187
Db 132 QKFVDK-PKEIAPTFLVQSKDDYTTSTKANIESLASAFGTGTQAKKEELAKLQKSIQ 190
QY 188 Y-----KKEITMDKNQKVLPAVAAKSGLLAHPNSVVGQFLSQ-LGFKEALSDDVTK 238
Db 191 VATKNESDKKALAILLNGKMAAFGAKSRF-----SFLVQTLKFKPT---DITF 237
QY 239 GLSKYLKGYLQMTTSLQVNPFRPIMTNKA-----SSNEPSLKE--LEKDPVWKKL 290
Db 238 EDSRHQ-----EVSPEVSKEINPDILFVINRITLAIGDSSNGVLENALIAETP----- 288
QY 291 NAVQNRVILDRDLWARSGLISSEMAKELVELSKK 328
Db 289 -AAKNGKIIQLTPDLVYLGGLGGLSTKLMIEDIQALK 325
RESULT 19
US-09-602-787A-598
; Sequence 598, Application US/09602787A
; Patent No. 6696561
; GENERAL INFORMATION:
; APPLICANT: Pompeius, Mark
; APPLICANT: Krüger, Burkhard
; APPLICANT: Schöler, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
```

```

; TITLE OF INVENTION: TRANSPORT
; FILE REFERENCE: BGI-125CP
; CURRENT APPLICATION NUMBER: US/09/602,787A
; CURRENT FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: USSN 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931454.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931478.0
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931563.9
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932122.1
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932124.8
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932180.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932182.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932190.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932191.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932209.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932212.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932227.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932228.7
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932229.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932230.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932927.3
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933005.0
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933006.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19940764.9
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940765.7
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940766.5
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940830.0
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940831.9
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940832.7
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940833.5
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19941378.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941395.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19942077.7
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942078.5
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942079.3
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942088.2
```

```
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 678
; SEQ ID NO 598
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-602-787A-598

Query Match      14.6%; Score 246.5; DB 4; Length 332;
Best Local Similarity 26.7%; Pred. No. 2.2e-12;
Matches 94; Conservative 70; Mismatches 129; Indels 59; Gaps 16;

Qy 9 KIL--SVIGLLFVLIATAACGNS-----SSNSSKSSKDGVEIKHEEGTTPVKPKRQV 62
Db 7 KILSSVVGVA-VLALLAGCSNADDTTADSTGTGNSAPFVSIEHEFTTIDVPERV 65

Qy 63 VLESHFDVALVALDVKPGVIADNKNRIIPLRDKI--GK-YTSVGRKOPNLEEISKL 119
Db 66 TLGVTDADIVLAGTVPVGTGYKFFENGLGFWTDELVEGKELTLLDSDSTPDLEQVAAL 125

Qy 120 KPDLIIADNNRHKG-IYKDLANKIAPTIELKSPFDGYN-ENIDAFKTIKALGKEEGKKR 177
Db 126 EPDLIIIGVSAGFDDVVEQLSDIAPVAVPAPAGTAAYAVAREEATNLVARAMQSEKGOEL 185

Qy 178 LEEHDKKIEEYKKEITMDKNQKVLPAVAAKSGLLAHP-----SNSVY-----GQFLSQLGF 228
Db 186 NEETDALIQAAAR-----DEN-----PSFDGKTGTVLIPYQGYGAYLPGDARGQFLDSLGI 236

Qy 229 ---KEALSDDDVTYKGLSKYLKPGYQLQMNTEITLSQVNPFRMFMTNKAASSNEPSLKELEKDP 285
Db 237 SLPEAVLSRDTGDSF-----FVDVPAESVKVDGVDVLLVL-----SNDENLDITAEANP 284

Qy 286 VWKXKLVNAGVQVDTILDRDLWARSGLI-----SSSEMAKELVELSK 327
Db 285 LPETLVNQQDAVIVATTE-----ERGAITYNSVLSPFALEHLAPRIAEALK 332

RESULT 20
US-09-328-352-4444
; Sequence 4444, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4444
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4444

Query Match      14.3%; Score 241; DB 4; Length 324;
Best Local Similarity 24.0%; Pred. No. 6e-12;
Matches 79; Conservative 77; Mismatches 135; Indels 38; Gaps 10;

Qy 14 IGLLFVLIATAACGNSSSNSSKSSK--DGVEIKHEEGTTPVKPKRQVVLVYSFVDA 71
Db 12 VALIIAAAVTQACQKQVADTTQASQKLAEPITVKHAGLTVIDHLPORVAVLNNEADF 71

Qy 72 LVALDVKPGVIADNKNRIIPLRDKIKGYTSVGRKOPNLEEISKLKPDIIADNNRH 131
Db 72 LDQLNVPIMGPKDVVPHFLEKYKD--AQIQDLGALVQPNMERIYALKPDLILM-TPLH 128

Qy 132 KGIYKDLANKIAPTIELKSPFDGYNEN-----IDAFKTIKALGKEEGKKRLEEHDK 183
Db 129 VNQYQLSKIAPTIH---YDINFNNSNHNGLVKDMMTGLGKIFNKEDLARQKVELDE 185

Qy 184 KIEEYKKEITMDKNQKVLPAVAAKSGLLAHPSNVYQFLSQLGPKEA--LSDDVTYKGLS 241
```

```
Db 186 QVKQV-QAVTANRPERALVVLHNNGAFSNFGIQSGYGFIFNAGFKVPASGVVDTSLHG-- 242
Qy 242 KYLKGPLYQMNTEITLSQVNPFRMFI-----MTWKASSNEPSLKELEKDPVWKLVNAVK 295
Db 243 -----QPISSEFIKKADPDILYIVDRTAVMEHRSNINAASV-----ENPLLRTQKAWKN 291
Qy 296 QRVDILDRDLWARSGLISSEEMAKELVE 324
Db 292 GRVIFVDADAWYTTAASTSLKIVWEDVK 320

RESULT 21
US-09-071-035-70
; Sequence 70, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 312 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-071-035-70

Query Match      14.2%; Score 240; DB 4; Length 312;
Best Local Similarity 27.3%; Pred. No. 6.9e-12;
Matches 98; Conservative 66; Mismatches 103; Indels 92; Gaps 20;

Qy 8 LKILSVIG--LLFVLIATAACGN-NSSNSSKSSKDGVEIKHEEGTT-----KVPKHPK 59
Db 4 LKKTVLIGTTLGSLFLLAACGNTNKEANNADKT-----HEVTDTLGNKVTVPAPKPK 55

Qy 60 RVVLEYSFV-DALVALDVKPVG---TADDNKKRIIKPLRDKIKGYTSVGRKQPNL-- 113
Db 56 RIIA---SYLSDYLVALGEKFPVQMTVGQSGIQDYLAKELKD-----VPTISY 100

Qy 114 ---EESKLPDLIIADNNR--HKGIYKDLANKIAPTIELKSPFDGYNENI---DAFKTI 164
Db 101 DLPEAVLVKFPDILLISSALVEGKGYKYSKIAPTIVVKN-----GENTVTRDQLEDI 155

Qy 165 SKALGKEEGKKRLEEHD---KKIIEY--KKEITMDKNQKVLPAVAAKSGLLAHPSNVY 219
Db 156 ATVLDDKKKQAKVLEDVDTLTKGVQVYLGKKD-----AGKSAAVLWVTNNQV 202
```

QY 220 -----GQFLSQ-LGFKEALSDDVTGKLSKYLKGPYLOMNTTTLQVNPFRMFMWN 269
Db 203 FMSVNRSGTLYQDLGLQ-----VPKVEEISKNATADQNVSLKLAELADHIFVNS 259
QY 270 KASSNEPSLKELEKDPVMKKLNAVKNQVRDILDRDLWARSRLISSEEMAKELVELSKK 328
Db 260 DESA-----PLFQEAIWKNLPAVKNNQVHTYDK-----KSSWLYNGPIANTQIVEDVKK 308

RESULT 22

US-08-961-083-24
; Sequence 24, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-961-083-24

Query Match 14.0%; Score 236; DB 3; Length 298;
Best Local Similarity 26.6%; Pred. No. 1.4e-11;
Matches 88; Conservative 65; Mismatches 112; Indels 66; Gaps 14;
QY 29 NSSNSSKSSKDGVB--IKHEGTTKVPKPKRVVVLVYSFVDALVALDVKPVGIADDN 86
Db 3 NSSTQSTSSAPTETVTKSSLDVKSVPKVIPTDGLGAADTIRALGF----- 53
QY 87 KKNRIKP-----LRDKIGKVTSGTRKQPNLEISKLKPDLLIADNNRHHGIYKDLN 139
Db 54 EKNVGMPTKTVPYTKDLVGVTKVNGSMKEPDLAIAALEPDLIIASPTQKFVDK-PK 112
QY 140 KIAPTIELKDFGDYNNENIDA-FKTIKALGK--BEEGKRLEHDKKIEEY----- 188
Db 113 EIAPTVLFQASKDDYWTSTKANIESLASAFGETGTQKAKEELTKLDSIQEVATKNESD 172
QY 189 KKEITMDKNQKVLPAVAKSGLLAHPNSYVVGFLSQ-LGFK--EALSDDVTKGLSKYLK 245
Db 173 KKAIALILLNEGKWAFAFGAKSRF-----SFLYQTLKFKPTDTKFDSDSRHG----- 216

QY 246 GPYLOMNTTTLQVNPFRMFMWNK-----SNBPSLKE--LEKDPVWKKLNAVKNQR 297
Db 217 ---QEVSFESVKEINPDILFVINRTLAIGGDSNNDGVLENALIAETP-----AAKNGK 267
QY 298 VDILDRDLWARSRLISSEEMAKELVELSKK 328
Db 268 IIQTPDLWYLSGGGLESTKLMIWEDIQALK 298

RESULT 23

US-09-536-784-24
; Sequence 24, Application US/09536784
; Patent No. 6573082
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/536,784
; FILING DATE: 30-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: OCT-30-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB340P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 24:
; US-09-536-784-24

Query Match 14.0%; Score 236; DB 4; Length 298;
Best Local Similarity 26.6%; Pred. No. 1.4e-11;
Matches 88; Conservative 65; Mismatches 112; Indels 66; Gaps 14;
QY 29 NSSNSSKSSKDGVB--IKHEGTTKVPKPKRVVVLVYSFVDALVALDVKPVGIADDN 86
Db 3 NSSTQSTSSAPTETVTKSSLDVKSVPKVIPTDGLGAADTIRALGF----- 53
QY 87 KKNRIKP-----LRDKIGKVTSGTRKQPNLEISKLKPDLLIADNNRHHGIYKDLN 139
Db 54 EKNVGMPTKTVPYTKDLVGVTKVNGSMKEPDLAIAALEPDLIIASPTQKFVDK-PK 112
QY 140 KIAPTIELKDFGDYNNENIDA-FKTIKALGK--BEEGKRLEHDKKIEEY----- 188
Db 113 EIAPTVLFQASKDDYWTSTKANIESLASAFGETGTQKAKEELTKLDSIQEVATKNESD 172
QY 189 KKEITMDKNQKVLPAVAKSGLLAHPNSYVVGFLSQ-LGFK--EALSDDVTKGLSKYLK 245
Db 173 KKAIALILLNEGKWAFAFGAKSRF-----SFLYQTLKFKPTDTKFDSDSRHG----- 216

131 HKGIYKOLNKIAPTIELKSFQDYNEN-----IDAPKTIKALGKEEGKRLKEEHDKKIE 186
128 AKAPDK-LNEIAPTIENTADTANLAKESAKERIDA---LAQIFGKQAEADKLKAEIDASPE 183
187 EYKKEITMDKNQKVLPAVAAGSGLLAHPSNSYVGFQFSLQGFKEALSDDVTGKLSKYLKG 246
184 A-AKTAAGKQKGLVILVNGGKMSAFGPSRLGWLHKDIGV-PAVDESIKEGSHGQ---- 238
247 PYLQMTTETLSQVNPFRMFMINTKASSNE--PSLKELEKDPVWKKLNAVKNORVILDRD 304
239 ---PISFYLKKNPDLFVLDRAAIGEGQAADKVLNPLVAETTTAKKGQVVYLVPE 295
305 LWARSRLGISSEEMAKELVELSKK 328
296 TYLAAGG-----AQELLNASKQ 312

RESULT 28
US-09-107-433-4519
; Sequence 4519, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 4519:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 295 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...295
; SEQUENCE DESCRIPTION: SEQ ID NO: 4519:
US-09-107-433-4519

Query Match 12.4%; Score 209; DB 4; Length 295;
Best Local Similarity 28.0%; Pred. No. 2.3e-09;
Matches 85; Conservative 53; Mismatches 110; Indels 56; Gaps 15;

131 HKGIYKOLNKIAPTIELKSFQDYNEN-----IDAPKTIKALGKEEGKRLKEEHDKKIE 186
128 AKAPDK-LNEIAPTIENTADTANLAKESAKERIDA---LAQIFGKQAEADKLKAEIDASPE 183
187 EYKKEITMDKNQKVLPAVAAGSGLLAHPSNSYVGFQFSLQGFKEALSDDVTGKLSKYLKG 246
184 A-AKTAAGKQKGLVILVNGGKMSAFGPSRLGWLHKDIGV-PAVDESIKEGSHGQ---- 238
247 PYLQMTTETLSQVNPFRMFMINTKASSNE--PSLKELEKDPVWKKLNAVKNORVILDRD 304
239 ---PISFYLKKNPDLFVLDRAAIGEGQAADKVLNPLVAETTTAKKGQVVYLVPE 295
305 LWARSRLGISSEEMAKELVELSKK 328
296 TYLAAGG-----AQELLNASKQ 312

RESULT 27
US-09-302-626B-10
; Sequence 10, Application US/09302626B
; Patent No. 6709660
; GENERAL INFORMATION:
; APPLICANT: Scalato, Enzo
; APPLICANT: Masignani, Vega
; APPLICANT: Rappuoli, Rino
; APPLICANT: Pizza, Mariagrazia
; APPLICANT: Grandi, Guido
; TITLE OF INVENTION: Meningococcal Antigens
; FILE REFERENCE: CHIR0159
; CURRENT APPLICATION NUMBER: US/09/302,626B
; CURRENT FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: PCT/IB99/00103
; PRIOR FILING DATE: 1999-01-14
; NUMBER OF SEQ ID NOS: 195
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 10
; LENGTH: 321
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-302-626B-10

Query Match 12.8%; Score 217; DB 4; Length 321;
Best Local Similarity 26.2%; Pred. No. 5.7e-10;
Matches 85; Conservative 63; Mismatches 136; Indels 40; Gaps 14;

21 IATAAGC--NNSSNSSKSSK-----GVEIKHEEGTTPKPKRNVVLEYSFVDA 71
13 LALGACSPQNSAPQAEQVSAQSEGVSVTVKTARGDVQIPQNPRIAYVDLGLMDT 72
72 LVALDVKPVGIIADNKNRIIPKLDKIGKITS-VGTRKQPNLEISKLKPOLIIADNRR 130
73 LSKLGVK-TGLSVD--KNRL--PYLEEFKTKPAGTLPEDPYETLNAYKPOLIIIGSRA 127
131 HKGIYKOLNKIAPTIELKSFQDYNEN-----IDAPKTIKALGKEEGKRLKEEHDKKIE 186
128 AKAPDK-LNEIAPTIENTADTANLAKESAKERIDA---LAQIFGKQAEADKLKAEIDASPE 183
187 EYKKEITMDKNQKVLPAVAAGSGLLAHPSNSYVGFQFSLQGFKEALSDDVTGKLSKYLKG 246
184 A-AKTAAGKQKGLVILVNGGKMSAFGPSRLGWLHKDIGVPAV--DEAIK-----EG 234
247 PYLQ-MNTETLSQVNPFRMFMINTKASSNE--PSLKELEKDPVWKKLNAVKNORVILDR 303
235 SHGQPSIFPYLKEKPNPDLFVLDRAAIGEGQAADKVLNPLVAETTTAKKGQVVYLVPE 294
304 DLWARSRLGISSEEMAKELVELSKK 328
295 ETYLAAGG-----AQELLNASKQ 312

MEDIUM TYPE: CD/ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 5819:
SEQUENCE CHARACTERISTICS:
LENGTH: 338 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...338
SEQUENCE DESCRIPTION: SEQ ID NO: 5819:
US-09-107-532A-5819

Query Match 11.5%; Score 194.5; DB 4; Length 338;
Best Local Similarity 23.4%; Pred. No. 4.4e-08;
Matches 78; Conservative 69; Mismatches 107; Indels 79; Gaps 16;
QY 23 TAACGNNSSNSKESKGGVEIKHEEGTKVPHKRVV--VLEYSFVDALVALDKVP 80
DB 51 TATTNSTSTSTHTVDTLGHK-----VDIPNAPKRIIGSYLE----NYLISLGERPI 101
QY 81 G---IADDNKNRIIPLRDKIGKTSVGRKQPNL-----EISKLKPLDIIADNNR- 130
DB 102 AQMTVGSGSIQYLDQLKD-----VPTISYDLPYEKVLSPFDPDLLLSSAT 149
QY 131 -HKGIVKDLNKIAPTIELKS-FPGDYNENIDAFKTSKALGKEEGKRLLEHDKKIEBY 188
DB 150 VEGGKTEQYSKIAPTVVKNGTDTWEEQL---KDVGKALAKEAEKVISDYQKSVST 206
QY 189 KKEITMDKNQKVLPAVA-----AKSGLLAHPNSVYVGFSLQGLGPKALSD 235
DB 207 REELA-DKIENKTAVALVWTNNSAFVSKNRSSGIIVDDLKFGVPDLVEEVSKEATSD- 264
QY 236 VTGKLSKYLKGPYLQWNTETLSQVNERFMFT-----NKASSNEPSLKELEKDPVKKLNA 292
DB 265 -----WSAVSSEKLAELADADYIILVNSDKDAAMFNPS-----WKNLKA 303
QY 293 VKQNRV-DILDRDLWARSRLGISSEEMAKELVE 324
DB 304 VKDNHVLFQFPESWLYN-GPIASQEMVNDIKE 335

FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 9745
LENGTH: 316
TYPE: PRT
ORGANISM: Myxococcus xanthus
US-09-902-540-9745

Query Match 10.9%; Score 185; DB 4; Length 316;
Best Local Similarity 22.8%; Pred. No. 2.4e-07;
Matches 64; Conservative 51; Mismatches 136; Indels 30; Gaps 9;
QY 44 EKHHEGTTKVPKPKRVVLEYSFVDALVALDKVPKGIADNNKRIIPLRDKIGKYT 103
DB 25 ELKHETGTLSPATPSRLAVYDGLVDTLNLGIRAVAVPK-----ATFPATLSAYN 76
QY 104 S-----VGRKQPNLEISKLKPLDIIADNNRHHKGIYKDLNKIAPTIELKSPDGDYNNI 158
DB 77 SNGVPOAGSLFEPDTKALAEIKPLDIIVVGSRKQA-ESLSNALAPTVNMSVRADHFVEDV 135
QY 159 DA-FKTIKALGKEEGKRLLEHDKKIEYKKEITMDKNQKVLPAVAAKSGLLAHPNS 217
DB 136 TGQTLALGKAFGQTEADTLVKNFEKERAALKK-LTGKSAVML--FTLNGNVVHVHAPGD 192
QY 218 YVGQFLSOLGPKAL--SDDVTGKLSKYLKGPYLQWNTETLSQV-----NPERFMFT 268
DB 193 RFGYLHETLGFNAVVPKAEAGSNDGPRPPAGSPKARQEAALAKKAAVAAPEWLVLVD 252
QY 269 NKASS--NEPSLKELEKDPVKKLNAVKNORVDILDRDLW 306
DB 253 RGAATGGGKNALETILAKHPELSQTRAFQKGRVIVVDAPSW 293

RESULT 33
US-09-302-626B-159
Sequence 159, Application US/09302626B
Patent No. 6709660
GENERAL INFORMATION:
APPLICANT: Scalato, Enzo
APPLICANT: Masignani, Vega
APPLICANT: Rappuoli, Rino
APPLICANT: Pizza, Mariagrazia
TITLE OF INVENTION: Meningococcal Antigens
FILE REFERENCE: CHIR0159
CURRENT APPLICATION NUMBER: US/09/302,626B
CURRENT FILING DATE: 1999-04-30
PRIOR APPLICATION NUMBER: PCT/IB99/00103
PRIOR FILING DATE: 1999-01-14
NUMBER OF SEQ ID NOS: 195
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 159
LENGTH: 240
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: ORF38a
US-09-302-626B-159

Query Match 10.7%; Score 181.5; DB 4; Length 240;
Best Local Similarity 32.2%; Pred. No. 3.2e-07;
Matches 59; Conservative 34; Mismatches 67; Indels 23; Gaps 10;
QY 21 IATAAG--NNSSNSSSKE-----SSKQGE--IKHEEGTKVPHKPKRVVLEYSFVDA 71
DB 13 LAIGACSPQNSDSPAQAEQAVSAQSEGSVTVTKTARGDVQIQPQFERIAVYDLGMLDT 72
QY 72 LVALDVKPVGIADNNKRIIPLRDKIGKITS-VGRKQPNLEISKLKPLDIIADNNR 130
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

```

Db      73  LSKLGVK-TGLSVD--KNRL--PYLEEFKTKKAGTLFEPDYETLNAYKQQLIIIGRA 127
Qy      131 HKGIYKDLNKIAPTIELKSPGDYDYN-----IDAFKTIISKALGKEEKGKRLBHDKKIE 186
Db      128 AKAFDK-LNEIAPTIENTADTNLKSNAKERIDA---LAQIFGKKAADKLKAEIDASF 183
Qy      187 EYK 189
Db      184 AAK 186

RESULT 34
US-09-489-039A-13095
; Sequence 13095, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709, 2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13095
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13095

```

```

; CURRENT FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: PCT/IB99/00103
; PRIOR FILING DATE: 1999-01-14
; NUMBER OF SEQ ID NOS: 195
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (142)..(143)
; OTHER INFORMATION: unknown
;
US-09-302-626B-8

Query Match          9.1%; Score 153; DB 4; Length 164;
Best Local Similarity 30.1%; Pred. No. 4.3e-05;
Matches 47; Conservative 31; Mismatches 62; Indels 16; Gaps 7;

QY  21  IATAACG--NNSSNSKSKESKD-----GVEIKHEEGTTKVPKPKRVVVVLEYSFVDA 71
DB  13  LALGACSPQNSDAPQAQKQAVSAAQTEGASVTVKTARGDVQIQPQPERIAVYDLGMLDT 72
QY  72  LVALDVKPGVIADDNKNRIKPLRDKIGKVTs-VGTRKQPNLEEISKLKPDLLIADNNR 130
DB  73  LSLKGVK-TGLSVD--KNRL--PYLEEYFKTKPAGTLFEPDYETINAYKPOLIIGSRA 127
QY  131  HKGIYKDLNKIAPTIELKSPDGDYNNENIDAPKTIK 166
DB  128  AKAFDK-LNEIATPIXTATNLKESAKEASTLAQ 162

RESULT 36
US-09-489-039A-7331
; Sequence 7311, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7331
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7331

Query Match          9.0%; Score 152; DB 4; Length 397;
Best Local Similarity 22.3%; Pred. No. 0.00017;
Matches 69; Conservative 61; Mismatches 119; Indels 60; Gaps 13;

QY  54  VPXHPKRVVVLEYSFVDALVALDVKP--VGIADDNKNRIIKP---LRDKIGKYTSVGT 107
DB  83  LPDHTRVVVVLQHTQLLLVQLHAAEDIVGVLSWQKQ--LGPQFARPMPEIGQLATPGD 140
QY  108  RKQPNLEEISKLKPDL-----IIADNNRHGIYKDLNKIAPTII- 145
DB  141  LTQVNIESLLARPPQVVFVANYAPPAMIAIQQAGIPVAISLRHDA-GEKNKNPTMA 199
QY  146  -ELKSPDGDYNNENIDAPKTIKSKALGKEEGKKRLLEHDKKIEEYKKEIT-----MDKNQK 199
DB  200  DEEQVYNAGLMLEGI-----RLTIGEVVERQPEAEALHYTFEARKQANAFVADIPPNQR 252
QY  200  VLPVAAKSGLLAHPSNYSVQGFSLQSGKEALSDDDVTGLSKYLVKGPYLMQNTTLLSQV 259
DB  253  V-RVYMANPDNTYGAGKYTGLMMAHAGALNVAASV-KGAR-----QVSLQGVLEW 302
QY  260  NPERMFWTNTKASNSNEPSLKELEKDPVWKLNNAVKNQRVDILDRDLWARSGLISSEEMA 319

```

```
Db 303 NPOVIFQDRYPQ-----VVKQIENDPQQAIDAVKHHRVWLMP-----YAKAWGYPMPPEALA 356
Qy 320 KELVELSKK 328
Db 357 LGELWMAKK 365

RESULT 37
US-09-489-039A-12270
; Sequence 12270, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: 2709.2004001
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12270
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12270

Query Match      8.5%; Score 144; DB 4; Length 353;
Best Local Similarity 20.9%; Pred. No. 0.00067;
Matches 67; Conservative 52; Mismatches 137; Indels 64; Gaps 10;

Qy 26 CGNSSSSSSKSSKSGDGVKIEKHEEGTKVPKH-----
Db 19 CPDAGGTAQRHS--GTDLRHPDGDPAIPGCGAGKCLLMMNPTLITRRLLIAMALS 76
Qy 58 -----PKRVVVLSEYFVDALVDKVPVGIAD-DNKKNRIRKP-LRDKIG 100
Db 77 PLLQMRGAQAADVDPQVVALEWLPAAELLALGVTVPYGVADIPNVLWVNEPALPDSV- 135
Qy 101 KYTSVTRKQPNLEISKLPDLIIADNNRHHGIYKDLNKPAPTEILKSPDGDYENIDA 160
Db 136 --IDVGLRTEPNLELLTKMPGFIVMSAGYGPSPEK-LARIAPGRGFTSPDGKRPLAMAQ 192
Qy 161 PFTISA--LCGEEGKKLEEDHKIEYKKEITMDKNQKVP-AAAAGSLLAHPSNS 217
Db 193 RSLLEMADLLGKTQQAQRHLAEFDALMESLRPRFAGRGDRPLMLTSLDPRHVLVFGNC 252
Qy 218 YVQFSLQGLFKEALSDVTKGLSKYLKGPYLMQNTETLSQVNPFRMFTMTNKASSNEPS 277
Db 253 LFQEVLDRFGIKNWTHGEAAFWGS-----VSVGIDRLAAFNADVICFDH---GNERD 302
Qy 278 LKELEKDPVWKKLNAVKNQR 297
Db 303 MAQLLATPLWQAMPFVRAGR 322

RESULT 38
US-09-489-039A-7324
; Sequence 7324, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: 2709.2004001
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7324
; LENGTH: 369
```

```
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7324

Query Match      8.5%; Score 143.5; DB 4; Length 369;
Best Local Similarity 21.9%; Pred. No. 0.00079;
Matches 56; Conservative 47; Mismatches 90; Indels 63; Gaps 11;

Qy 51 TTKVPKPKRVVVL-EYSFVDALVALDVKVP-----GTAADNKK 88
Db 84 TTVPLHPQIRVSMHLDITITPLIELGAPPIASHGRTRPDGSHYLRSSAQLTGVDFDNSD 143
Qy 89 NRIIKPLRDKTKYTSVTRKQPNLEISKLPDLIIADNNRHHGIYKDLNKPAPTEILK 148
Db 144 IRFI-----GT-ADIDLEAVAAARPLIIITEPSRHVSF-EQLSKIAPTVSID 188
Qy 149 SFGDYNENIDAFKTIISKALKEEGKKRLEEDHKIEYKKEITMDKN--QKVLPAVAA 206
Db 189 HLQSAFE---LYRKLALQLTGTQ-----PRLAILERRYQEQIKQLKAMVNPQYSVSVIOA 241
Qy 207 KSG-LLAHPNSNYVQFSLQGLFK-----EALSDDVTKGLSKYLKGPYLMQNTETLSQVN 260
Db 242 NNGKVTVHHSYHALGRVLRDAGFRPPPLIERIPD-----GQRIDVSAEQLELPD 290
Qy 261 PERFMINTKASSNEP 276
Db 291 ADFVFATWRSDTGGKP 306

RESULT 39
US-09-328-352-7262
; Sequence 7262, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7262
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7262

Query Match      8.3%; Score 140.5; DB 4; Length 354;
Best Local Similarity 22.9%; Pred. No. 0.0013;
Matches 75; Conservative 56; Mismatches 113; Indels 83; Gaps 15;

Qy 10 ILSVIGLLFVLATAACGNSSSSKSSKSGDGVKIEKHEEGTK--VPKPKRVVVLSEYS 67
Db 15 IMKLASLTLTLLICGSSSEONRKTQENSTSHSCVFDSTNTKVCVAKPAQRIYSLPES 74
Qy 68 FVDALVALDV--KPVGIADD-----NKKNRIRIKPLRDKIKGKYTSVTRKQPNLEIS 117
Db 75 GLDGLYMLGQGDKVIGIPAEVVIQPLLFNAYSIDKRIANKQLAAPSQGA-NATNIESIV 133
Qy 118 KLKPLDIIADNNRHHGIYKDLNKPAPTEILKSP-----DGYNENIDAFKTIISKAL 168
Db 134 LLKPLDLIVVSGQ-----TQTIELLRFQGIADVVMBSGTYKQVKESELSEIALIS 182
Qy 169 GKEEGSKRLEEDHKIEYKKEITMDKNQKLPVAAAKSGLLAHPSNSYV---GQFLS 224
Db 183 GAQKAQAQILNPSDEIVAE-----VAAKTARQPNKQSIYYAWSGGRIFS 226
Qy 225 QLGKFEALSDVTKGLSKYLKGPY-----LQMTTETLSQVNPFRMFTMTNKASSNE 275
Db 227 TSG-RESITNDFIE-----LAGAYNIQTNANQPNVNPETLIEWNPDLVLM-----NTN 275
Qy 276 PSL-----KELEKDPVWKKLNAVKNQR 298
Db 276 PSL-----KELEKDPVWKKLNAVKNQR 298
```

Db 276 PKLIYERKELQ-----GLSAVQNRKV 296

RESULT 40

US-08-961-083-194
; Sequence 194, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961.083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 194:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 132 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-961-083-194

Query Match 7.9%; Score 134; DB 3; Length 132;
Best Local Similarity 30.5%; Pred No. 0.0012;
Matches 39; Conservative 23; Mismatches 52; Indels 14; Gaps 4;
Qy 21 IATAACGNNS---SSNSKSKSGKGVEIKHEEGTTKPKPKRVVVLVYSFVDALVALDV 77
Db 2 IVFSACSSNSKXNEENTSKEHAPDKIVLDHAFGGQITLDKPKERVATIAWGNHHDVALALGI 61
Qy 78 KPVGIADN---KKRIKIP-----LRDKIGKYSVGTGRKQPNLEEISKLKPDLLIADNN 129
Db 62 VPVGFSKANYGVSDKGVLPWTEEKIKELNGKANLFDLGLNFEAISNSKPDVILAG-- 119
Qy 130 RHKGIYKD 137
Db 120 -YSGITKE 126

Search completed: August 25, 2005, 09:27:38
Job time : 26 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 09:17:46 ; Search time 65 seconds
(without alignments)

2639.177 Million cell updates/sec

Title: US-10-724-972A-6352

Perfect score: 1690

Sequence: 1 GVESVRLKILSVIGLLFVL.....EEMAKVELSKDSKDNK 335

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1681	99.5	334	2	Q8CNG3	Q8CNG3 staphylococ
2	1232	72.9	327	2	Q99S89	Q99S89 staphylococ
3	1232	72.9	327	2	Q7A499	Q7A499 staphylococ
4	1230	72.8	327	2	Q6GQK2	Q6GQK2 staphylococ
5	1227	72.6	327	2	Q8NVD6	Q8NVD6 staphylococ
6	1227	72.6	327	2	Q6G7D7	Q6G7D7 staphylococ
7	725.5	42.9	319	2	Q8NXA6	Q8NXA6 staphylococ
8	725.5	42.9	319	2	Q6GAH5	Q6GAH5 staphylococ
9	724.5	42.9	319	2	Q99V55	Q99V55 staphylococ
10	724.5	42.9	319	2	Q7A6B6	Q7A6B6 staphylococ
11	719.5	42.6	319	2	Q6G145	Q6G145 staphylococ
12	679.5	40.2	315	2	Q34348	Q34348 bacillus su
13	611	36.2	348	2	O07616	O07616 bacillus su
14	608.5	36.0	325	2	Q65LP5	Q65LP5 bacillus li
15	516.5	30.6	298	2	Q9C8B7	Q9C8B7 pasteurilla
16	488.5	28.9	307	2	Q8GRB3	Q8GRB3 vibrio para
17	487.5	28.8	307	2	Q87FM4	Q87FM4 vibrio para
18	476	28.2	302	2	Q7M254	Q7M254 photorhabdu
19	476	28.2	304	2	Q6D898	Q6D898 erwinia car
20	463	27.4	300	2	Q93F19	Q93F19 shigella fl
21	460	27.2	300	1	FECS_ECOLI	FECS_ECOLI
22	454.5	26.9	306	2	Q88916	Q88916 pseudomonas
23	415.5	24.6	325	2	Q65F62	Q65F62 bacillus li
24	405	24.0	330	2	Q87491	Q87491 staphylococ
25	405	24.0	330	2	Q7A127	Q7A127 staphylococ
26	405	24.0	330	2	Q7A2Y7	Q7A2Y7 staphylococ
27	405	24.0	330	2	Q7A869	Q7A869 staphylococ
28	405	24.0	330	2	Q6GD10	Q6GD10 staphylococ
29	405	24.0	330	2	Q6GKJ0	Q6GKJ0 staphylococ
30	404.5	23.9	321	2	Q81V85	Q81V85 bacillus an
31	404	23.9	321	2	Q6HNJ9	Q6HNJ9 bacillus th

32 399 23.6 321 2 Q63G31 Q63G31 bacillus ce
33 397 23.5 321 2 Q73DM9 Q73DM9 bacillus ce
34 396.5 23.5 322 2 Q81I09 Q81I09 bacillus ce
35 392 23.2 324 2 Q817L3 Q817L3 bacillus ce
36 391 23.1 320 2 Q81DY8 Q81DY8 bacillus ce
37 391 23.1 324 2 Q81L65 Q81L65 bacillus th
38 389 23.0 324 2 Q6HCZ7 Q6HCZ7 bacillus li
39 380 22.5 315 2 Q65NE2 Q65NE2 bacillus li
40 379.5 22.5 315 2 Q65EN6 Q65EN6 bacillus li
41 379.5 22.5 321 2 Q738W0 Q738W0 bacillus ce
42 378 22.4 316 2 Q8XM98 Q8XM98 clostridium
43 378 22.4 320 2 Q81R05 Q81R05 bacillus an
44 377 22.3 320 2 Q6HJA9 Q6HJA9 bacillus th
45 373 22.1 320 2 Q63BT9 Q63BT9 bacillus ce

ALIGNMENTS

RESULT 1
Q8CNG3 PRELIMINARY; PRT; 334 AA.
AC Q8CNG3;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Ferrichrome ABC transporter.
GN OrderedLocustNames=SEL768;
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12228;
RX PubMed=12950922;
RA Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J.,
Qin Z.-Q., Miao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,
Yuan Z.-H., Zhao G.-P., Qu D., Danchin A., Wen Y.-M.;
RT "Genome-based analysis of virulence genes in a non-biofilm-forming
RT Staphylococcus epidermidis strain (ATCC 12228).";
RL Mol. Microbiol. 49:1577-1593 (2003).
DR EMBL; AE016750; AA05409.1; -
DR GO; GO:0005381; F:iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2_1.
KW Complete proteome.
SQ SEQUENCE 334 AA; 37438 MW; 92C02D41A5ABAF33 CRC64;

Query Match 99.5%; Score 1681; DB 2; Length 334;
Best Local Similarity 99.7%; Pred. No. 5.2e-84;
Matches 333; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 VESVRLKILSVIGLLFVLIIATAACNNSSNSSKSSKSGDVEIKHEEGTTKVPKPKRV 61
DB 1 VESVRLKILSVIGLLFVLIIATAACNNSSNSSKSSKSGDVEIKHEEGTTKVPKPKRV 60
QY 62 VVLEYSFVDALVALDVKPGVIGADDNKKRIIPLRDKIGKYSVTGRKQPNLEEISKLKP 121
DB 61 VVLEYSFVDALVALDVKPGVIGADDNKKRIIPLRDKIGKYSVTGRKQPNLEEISKLKP 120
QY 122 DLIIADNRHKGIIYKDLNKAFTPIELKSPDGDYNNENIDAFKTIISKALGKEEGKKLEEH 181
DB 121 DLIIADNRHKGIIYKDLNKAFTPIELKSPDGDYNNENIDAFKTIISKALGKEEGKKLEEH 180
QY 182 DKKIEYKKEITMDKNQKVLPAVAAKGLLAHPSNVYQFLSQLGFKALSDDVTGKLS 241
DB 181 DKKIEYKKEITMDKNQKVLPAVAAKGLLAHPSNVYQFLSQLGFKALSDDVTGKLS 240
QY 242 KYLGKPYLQMTTETLSQVNPFRFMTNKAASSNEPSLKELEKDPVWKLVNAVKQNRVDIL 301
DB 241 KYLGKPYLQMTTETLSQVNPFRFMTNKAASSNEPSLKELEKDPVWKLVNAVKQNRVDIL 300

```

Qy 302 DRDLWARSRLGISSEEMAKELVELSKKDKXK 335
Db 301 DRDLWARSRLGISSEEMAKELVELSKKDKXK 334

RESULT 2
Q99889 PRELIMINARY; PRT; 327 AA.
ID Q99889
AC Q99889;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to ferrichrome ABC transporter.
GN OrderedLocusNames=SAV2177;
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Mu50 / ATCC 700699;
RX MEDLINE=21311952; PubMed=11418146; DOI=10.1016/S0140-6736(00)04403-2;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003364; BAB58339.1; -.
DR PIR; D90013; D90013.
DR GO; GO:0005381; P:iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2; 1.
KW Complete proteome.
SQ SEQUENCE 327 AA; 36591 MW; 6AEECBC4E17CDB7 CRC64;

Query Match 72.9%; Score 1232; DB 2; Length 327;
Best Local Similarity 71.9%; Pred. No. 1.5e-59;
Matches 235; Conservative 42; Mismatches 50; Indels 0; Gaps 0;

Qy 5 VRGLKILSVIGLLFVLVIAAAGNNSNSSKESKDGVEIKHBEETTKVPHKPRVVL 64
Db 1 MRGLKTFSLGLVALLVAACGNTDSSKESSTKDTISVKDENGTVKVPKDAKRIIVL 60

Qy 65 BYSFVDALVALDVKPGVGIADNKKRIIKPLDKIKGYTSVGTTRKQPNLEISLKPDLI 124
Db 61 EYSFADALALDVKPGVGIADNKKRIIKPVREKIGDYTSVGTTRKQPNLEISLKPDLI 120

Qy 125 IADNRHKGIVKDLNKIAPTIELKSFQGDYNNENIDAPKTIKALGKEEGKRLLEHDKK 184
Db 121 IADSRHKGINKELNKIAPTLSLKSFGDYKQNTNSFKTIKALNKEKEGKRLAEHDKL 180

Qy 185 IEYKKEITMDKNQKVLPAVAAKSGLLAHPNSYVGQFLSQGFKEALSDDDVTGSLSKYL 244
Db 181 INKYKDEIKPDRNQKVLPAVAKAGLLAHPNYSYVGQFLNELGPKNALSDDDVTGSLSKYL 240

Qy 245 KGPYLQMTETLSQVNPFRMFTMTNKASSNPSPSLKELEKOPVWKKLNAVKNQVRDILDRD 304
Db 241 KGPYLQLDTEHLADLNPRMIIMTDHAKKDSAEFKKLQEDATWKKLNAVKNRVDIVDRD 300

Qy 305 LWARSRLGISSEEMAKELVELSKKDKK 331
Db 301 VWARSRLGISSEEMAKELVELSKKQK 327

RESULT 3
Q7A499 PRELIMINARY; PRT; 327 AA.
ID Q7A499
AC Q7A499;

```

```

DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE SAI979 protein.
GN OrderedLocusNames=SAI979;
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158879;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21311952; PubMed=11418146; DOI=10.1016/S0140-6736(00)04403-2;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003364; BAB58339.1; -.
DR GO; GO:0005381; P:iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2; 1.
KW Complete proteome.
SQ SEQUENCE 327 AA; 36591 MW; 6AEECBC4E17CDB7 CRC64;

Query Match 72.9%; Score 1232; DB 2; Length 327;
Best Local Similarity 71.9%; Pred. No. 1.5e-59;
Matches 235; Conservative 42; Mismatches 50; Indels 0; Gaps 0;

Qy 5 VRGLKILSVIGLLFVLVIAAAGNNSNSSKESKDGVEIKHBEETTKVPHKPRVVL 64
Db 1 MRGLKTFSLGLVALLVAACGNTDSSKESSTKDTISVKDENGTVKVPKDAKRIIVL 60

Qy 65 BYSFVDALVALDVKPGVGIADNKKRIIKPLDKIKGYTSVGTTRKQPNLEISLKPDLI 124
Db 61 EYSFADALALDVKPGVGIADNKKRIIKPVREKIGDYTSVGTTRKQPNLEISLKPDLI 120

Qy 125 IADNRHKGIVKDLNKIAPTIELKSFQGDYNNENIDAPKTIKALGKEEGKRLLEHDKK 184
Db 121 IADSRHKGINKELNKIAPTLSLKSFGDYKQNTNSFKTIKALNKEKEGKRLAEHDKL 180

Qy 185 IEYKKEITMDKNQKVLPAVAAKSGLLAHPNSYVGQFLSQGFKEALSDDDVTGSLSKYL 244
Db 181 INKYKDEIKPDRNQKVLPAVAKAGLLAHPNYSYVGQFLNELGPKNALSDDDVTGSLSKYL 240

Qy 245 KGPYLQMTETLSQVNPFRMFTMTNKASSNPSPSLKELEKOPVWKKLNAVKNQVRDILDRD 304
Db 241 KGPYLQLDTEHLADLNPRMIIMTDHAKKDSAEFKKLQEDATWKKLNAVKNRVDIVDRD 300

Qy 305 LWARSRLGISSEEMAKELVELSKKDKK 331
Db 301 VWARSRLGISSEEMAKELVELSKKQK 327

RESULT 4
Q6GE02 PRELIMINARY; PRT; 327 AA.
ID Q6GE02
AC Q6GE02;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative transport system binding lipoprotein.
GN OrderedLocusNames=SAR2268;
OS Staphylococcus aureus (strain MRSA252).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=282458;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15213324; DOI=10.1073/pnas.04025211101;

```


RA Holden M.T.G., Feil E.J., Lindsey J.A., Peacock S.J., Day N.P.J.,
RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,
RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,
RA Feltwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagels K.,
RA James K.D., Lennard N., Line A., Mayes R., Moule S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitch E., Rutherford K.M., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,
RA Spratt B.G., Parkhill J.,
RT "Complete genomes of two clinical *Staphylococcus aureus* strains:
RT evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
DR EMBL; BX571856; CAG41246.1; -;
DR GO; GO:0005381; F:iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2; 1.
DR Complete proteome.
KW GO; GO:0005381; F:iron ion transporter activity; IEA.
SQ SEQUENCE 327 AA; 36687 MW; AC34D0F769D71265 CRC64;

Query Match 72.8%; Score 1230; DB 2; Length 327;
Best Local Similarity 72.2%; Pred. No. 1.9e-59;
Matches 236; Conservative 39; Mismatches 52; Indels 0; Gaps 0;

QY 5 VRGLKILSVIGLLFLVLIATAACGNSSNSSKSSKSGDVEIKHEGGTTKVPKPKRVVVL 64
DB 1 MRGLKTFSLIGLIVALLFLVAACGNTDSSKESSTKDTISVDENGTVKVPKDAKRIIVL 60

QY 65 EYSFVDALVALDVKPVGIADNDKNNRIIKPLRDKIGKTVSGTRKQPNLEISKLKPDLI 124
DB 61 EYSFADALVALDVKPVGIADNDKNNRIIKPVREKIDYTSVGTTRKQPNLEISKLKPDLI 120

QY 125 IADNRHKGIIYKDLANKIAPTIELKSPDGYNNENIDAFKTIKALGKEEGKRLAEHDKK 184
DB 121 IADSRHKGINKELNKLAPTLKSLKSPDGYKQNIINSFKTIKALNKEKEGKRLAEHDKL 180

QY 185 IEYKKEITMDKNQKVLPAVAKSGLLAHPNSYVQFLSGLQFKEALSDDDVTGKLSKYL 244
DB 181 INKYKDEIKFDNRQKVLPAVAKAGLLAHPNSYVQFLNELGFKNALSDDDVTGKLSKYL 240

QY 245 KGPYLQMTETLSQVNPMPFIMTNKASNPFLKELEKDPVWKLNNAVKNQVRDILDRD 304
DB 241 KGPYLQDTEHLADLNPERMIIMTDNAKDSAEFKLQEDATWTKLNNAVKNRVDIVDRD 300

QY 305 LWASRGLISSEEMAKELVELSKDSK 331
DB 301 VWASRGLISSEEMAKELVELSKKEQK 327

RESULT 5

Q8NV06 ID Q8NV06 PRELIMINARY; PRT; 327 AA.
AC Q8NV06;
DT 01-OCT-2002 (T:EMBLrel. 22, Created)
DT 01-OCT-2002 (T:EMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (T:EMBLrel. 26, Last annotation update)
DE W2103 protein.
GN OrderedLocusNames=WM2103;
OS *Staphylococcus aureus* (strain WM2).
OC Bacteria; Firmicutes; Bacillales; *Staphylococcus*.
OX NCBI_TaxID=196620;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WM2;
RX MEDLINE=22040717; PubMed=12044378; DOI=10.1016/S0140-6736(02)08713-5;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA";
RL Lancet 359:1819-1827(2002).
DR EMBL; AP004829; BAB95968.1; -;
DR GO; GO:0005381; F:iron ion transporter activity; IEA.

DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2; 1.
DR Complete proteome.
KW GO; GO:0005381; F:iron ion transporter activity; IEA.
SQ SEQUENCE 327 AA; 36577 MW; 6AEDCDBC4E169CB7 CRC64;

Query Match 72.6%; Score 1227; DB 2; Length 327;
Best Local Similarity 71.6%; Pred. No. 2.8e-59;
Matches 234; Conservative 42; Mismatches 51; Indels 0; Gaps 0;

QY 5 VRGLKILSVIGLLFLVLIATAACGNSSNSSKSSKSGDVEIKHEGGTTKVPKPKRVVVL 64
DB 1 MRGLKTFSLIGLIVALLFLVAACGNTDSSKESSTKDTISVDENGTVKVPKDAKRIIVL 60

QY 65 EYSFVDALVALDVKPVGIADNDKNNRIIKPLRDKIGKTVSGTRKQPNLEISKLKPDLI 124
DB 61 EYSFADALVALDVKPVGIADNDKNNRIIKPVREKIDYTSVGTTRKQPNLEISKLKPDLI 120

QY 125 IADNRHKGIIYKDLANKIAPTIELKSPDGYNNENIDAFKTIKALGKEEGKRLAEHDKK 184
DB 121 IADSRHKGINKELNKLAPTLKSLKSPDGYKQNIINSFKTIKALNKEKEGKRLAEHDKL 180

QY 185 IEYKKEITMDKNQKVLPAVAKSGLLAHPNSYVQFLSGLQFKEALSDDDVTGKLSKYL 244
DB 181 INKYKDEIKFDNRQKVLPAVAKAGLLAHPNSYVQFLNELGFKNALSDDDVTGKLSKYL 240

QY 245 KGPYLQMTETLSQVNPMPFIMTNKASNPFLKELEKDPVWKLNNAVKNQVRDILDRD 304
DB 241 KGPYLQDTEHLADLNPERMIIMTDNAKDSAEFKLQEDATWTKLNNAVKNRVDIVDRD 300

QY 305 LWASRGLISSEEMAKELVELSKDSK 331
DB 301 VWASRGLISSEEMAKELVELSKKEQK 327

RESULT 6

Q6G7D7 ID Q6G7D7 PRELIMINARY; PRT; 327 AA.
AC Q6G7D7;
DT 05-JUL-2004 (T:EMBLrel. 27, Created)
DT 05-JUL-2004 (T:EMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (T:EMBLrel. 27, Last annotation update)
DE Putative transport system binding lipoprotein.
GN OrderedLocusNames=SA52078;
OS *Staphylococcus aureus* (strain MSSA476).
OC Bacteria; Firmicutes; Bacillales; *Staphylococcus*.
OX NCBI_TaxID=282459;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15213324; DOI=10.1073/pnas.0402521101;
RA Holden M.T.G., Feil E.J., Lindsey J.A., Peacock S.J., Day N.P.J.,
RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,
RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,
RA Feltwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagels K.,
RA James K.D., Lennard N., Line A., Mayes R., Moule S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitch E., Rutherford K.M., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,
RA Spratt B.G., Parkhill J.;
RT "Complete genomes of two clinical *Staphylococcus aureus* strains:
RT evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
DR EMBL; BX571857; CAG43886.1; -;
DR GO; GO:0005381; F:iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2; 1.
DR Complete proteome; Lipoprotein.
KW GO; GO:0005381; F:iron ion transporter activity; IEA.
SQ SEQUENCE 327 AA; 36577 MW; 6AEDCDBC4E169CB7 CRC64;

Query Match 72.6%; Score 1227; DB 2; Length 327;
Best Local Similarity 71.6%; Pred. No. 2.8e-59;
Matches 234; Conservative 42; Mismatches 51; Indels 0; Gaps 0;

```
Qy 5 VRGKILSVIGLLFVLITATAACGNSSSSKSSKDGVEIKHEGTTKVPKPKRVVVL 64
Db 1 MRGLKTFILGLVALLVAACGNTDSSKSSSTKDTISVKDNGTVKVPKDAKRIIVL 60
Qy 65 EYSPVDALVALDVKPVGIADGNKNRIIKPVRDKIGKYSVTGTRKQPNLEISIKLPDLI 124
Db 61 EYSPADALVALDVKPVGIADGNKKRIIKPVRDKIGKYSVTGTRKQPNLEISIKLPDLI 120
Qy 125 IADNRHKGIIYKDLNKIAPTIELKSPDGYNENIDAFKTSKALGKEEGKKRLEEHDKK 184
Db 121 IADSSRHKGINKELNKIAPTIELKSPDGYKQINSPKTIKALNKEKEGKRLAEHDKL 180
Qy 185 IEEYKKEITMDKNQKVLPAVAAKSGLLAHPNSVVGQFLSQGLKEALSDDVTGLSKYL 244
Db 181 INKYKDEIKFDRNQKVLPAVAAKAGLLAHPNSVVGQFLNELGFKNALSDDDVTGLSKYL 240
Qy 245 KGPVLQMTETLSOVNPERMFTMTNKASSNEPSLKELEKDPVWKLNNAVKNQVRDILDRD 304
Db 241 KGPVLQDTEHLADINPERMFTMTDHAKDSAEFKLQSDATWKKLNNAVKNRVDIVDRD 300
Qy 305 LWARSRLISSEEMAKELVELSKDSK 331
Db 301 VWARSRLISSEEMAKELVELSKKEQK 327

RESULT 7
Q8NXXA6 PRELIMINARY; PRT; 319 AA.
ID Q8NXXA6
AC Q8NXXA6
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE MW0921 protein.
GN OrderedLocuNames=MM0921;
OS Staphylococcus aureus (strain MW2);
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MM2;
RX MEDLINE=22040717; PubMed=12044378; DOI=10.1016/S0140-6736(02)08713-5;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
acquired MRSA.";
RL Lancet 359:1819-1827(2002).
DR EMBL; AP004825; BAB94786.1; -.
DR GO; GO:0005381; F:iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2; 1.
KW Complete proteome.
SQ SEQUENCE 319 AA; 35552 MW; DB386C3BB2C60535 CRC64;

Query Match 42.9%; Score 725.5; DB 2; Length 319;
Best Local Similarity 46.9%; Pred. No. 5.7e-32;
Matches 149; Conservative 63; Mismatches 95; Indels 11; Gaps 4;

Qy 13 VIGLLFVLIATAACGNSSSSKSSKDGVEIKHEGTTKVPKPKRVVVLVYSFVDAL 72
Db 9 VVFMILVAVAGCGQKDT-----BEKTEMTTIKDELGTETKIKNPKRVVVLVYSFADYL 63
Qy 73 VALDVKPVGIADGNKNRIIKPLRDKIGKYSVTGTRKQPNLEISIKLPDLIIADNRHK 132
Db 64 AALDMKPVGIADGSSKNITKSVRDKIGAYESVGSRPQPNMEVSKLPDLIIADVSRHK 123
Qy 133 GIYKDLNKIAPTIELKSPDGYNENIDAFKTSKALGKEEGKKRLEEHDKKIEYKKEI 192
Db 124 KIKSLSKIAPTITMLVSGTGDYNANIETAFKTVAAKVGKEGKELKHKDLIABIRKI 183
Qy 193 TMDKNQKVLPAVAAKSGLLAHPNSVVGQFLSQGLKEALSDDVTGLSKYLKGPYLQWN 252
Db 184 EQSTLKSAPAFGISRAGMFINNEDTFMGQFLKMGIOPEVTKDTHVGERKGGPYIYNL 243
Qy 253 TETLSOVNPERMFTMTN-KASSNEPSLKELEKDP-VWKLNNAVKNQVRDILDRDLWARSR 310
Db 244 NEELANINPKVMILATDGTDKNRTK-----IDPAVWKS LKAVKDKNVYDVRNKLKSR 299
Qy 311 GLISSEEMAKELVELSKK 328
Db 300 GTIASESMASDEKIAEK 317
```

```
Db 184 EQSTLKSAPAFGISRAGMFINNEDTFMGQFLKMGIOPEVTKDTHVGERKGGPYIYNL 243
Qy 253 TETLSOVNPERMFTMTN-KASSNEPSLKELEKDP-VWKLNNAVKNQVRDILDRDLWARSR 310
Db 244 NEELANINPKVMILATDGTDKNRTK-----IDPAVWKS LKAVKDKNVYDVRNKLKSR 299
Qy 311 GLISSEEMAKELVELSKK 328
Db 300 GTIASESMASDEKIAEK 317

RESULT 8
Q6GAH5 PRELIMINARY; PRT; 319 AA.
ID Q6GAH5
AC Q6GAH5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Transport system extracellular binding lipoprotein.
DE OrderedLocuNames=SAS0973;
GN Staphylococcus aureus (strain MSSA476).
OS Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=282459;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15213324; DOI=10.1073/pnas.0402521101;
RA Hoiden M.T.G., Feil E.J., Lindsay J.A., Peacock S.J., Day N.P.J.,
RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,
RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,
RA Feltwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagels K.,
RA James K.D., Lennard N., Line A., Mayes R., Moule S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitch E., Rutherford K.M., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,
RA Spratt B.G., Parkhill J.;
RT "Complete genomes of two clinical Staphylococcus aureus strains:
evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
DR EMBL; BX571857; CAG42748.1; -.
DR GO; GO:0005381; F:iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2; 1.
KW Complete proteome; Lipoprotein.
SQ SEQUENCE 319 AA; 35552 MW; DB386C3BB2C60535 CRC64;

Query Match 42.9%; Score 725.5; DB 2; Length 319;
Best Local Similarity 46.9%; Pred. No. 5.7e-32;
Matches 149; Conservative 63; Mismatches 95; Indels 11; Gaps 4;

Qy 13 VIGLLFVLIATAACGNSSSSKSSKDGVEIKHEGTTKVPKPKRVVVLVYSFVDAL 72
Db 9 VVFMILVAVAGCGQKDT-----BEKTEMTTIKDELGTETKIKNPKRVVVLVYSFADYL 63
Qy 73 VALDVKPVGIADGNKNRIIKPLRDKIGKYSVTGTRKQPNLEISIKLPDLIIADNRHK 132
Db 64 AALDMKPVGIADGSSKNITKSVRDKIGAYESVGSRPQPNMEVSKLPDLIIADVSRHK 123
Qy 133 GIYKDLNKIAPTIELKSPDGYNENIDAFKTSKALGKEEGKKRLEEHDKKIEYKKEI 192
Db 124 KIKSLSKIAPTITMLVSGTGDYNANIETAFKTVAAKVGKEGKELKHKDLIABIRKI 183
Qy 193 TMDKNQKVLPAVAAKSGLLAHPNSVVGQFLSQGLKEALSDDVTGLSKYLKGPYLQWN 252
Db 184 EQSTLKSAPAFGISRAGMFINNEDTFMGQFLKMGIOPEVTKDTHVGERKGGPYIYNL 243
Qy 253 TETLSOVNPERMFTMTN-KASSNEPSLKELEKDP-VWKLNNAVKNQVRDILDRDLWARSR 310
Db 244 NEELANINPKVMILATDGTDKNRTK-----IDPAVWKS LKAVKDKNVYDVRNKLKSR 299
Qy 311 GLISSEEMAKELVELSKK 328
Db 300 GTIASESMASDEKIAEK 317
```

Db 300 GIIASESMAEDLEKIAEK 317

RESULT 9

Q99V55 PRELIMINARY; PRT; 319 AA.

AC Q99V55;

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-JUN-2001 (TrEMBLrel. 26, Last annotation update)

DE Similar to ferrichrome ABC transporter.

GN OrderedLocusNames=SAV1038;

OS Staphylococcus aureus (strain Mu50 / ATCC 700699).

OC Bacteria; Firmicutes; Bacillales; Staphylococcus.

OX NCBI_TaxID=158878;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Mu50 / ATCC 700699;

RX MEDLINE=21311952; PubMed=11418146; DOI=10.1016/S0140-6736(00)04403-2;

RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I., Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T., Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A., Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C., Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J., Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.

RT "Whole genome sequencing of methicillin-resistant Staphylococcus aureus."

RL Lancet 357:1225-1240(2001).

DR EMBL; AP003361; BAB57200.1; -.

DR PIR; E89872; E89872.

DR GO; GO:0005381; F:iron ion transporter activity; IEA.

DR GO; GO:0006827; P:high affinity iron ion transport; IEA.

DR InterPro; IPR002491; Peripla_BP.

DR Pfam; PF01497; Peripla_BP_2; 1.

KW Complete proteome.

SQ SEQUENCE 319 AA; 35566 MW; DE6D7C7AB4831574 CRC64;

Query Match 42.9%; Score 724.5; DB 2; Length 319;

Best Local Similarity 46.9%; Pred. No. 6.5e-32;

Matches 149; Conservative 63; Mismatches 95; Indels 11; Gaps 4;

QY 13 VIGLLFVLIIATAACGNSSSSSKSSKSDGVEIKHEEGTTKVPKPKRVVVLVLEYSFVDAL 72

Db 9 VVFMILLVAVAGCGQKDT-----EKEETMTIKDELGTEKIKKPKRVVVLVLEYSFADYL 63

QY 73 VALDVKPVGIADDNKKRRIKPLRDKIGKYSVGTGRKQNPNEISIKLKPDLIIADNNRHK 132

Db 64 AALDMKPVGIADGSGTKNITKSVDRKIGAYESVGSRPQNMVEISKLPDLIIADVSRHK 123

QY 133 GIYKDLNKIAPTIELKSPDGYNENIDAKTISKALGKEEGKRLKEHDKKIEEYKKEI 192

Db 124 KIKSELSKIAPTIMLVSGTGDYNANIEAFKTVAKAVGKEGKELKHKDLAEIRKKI 183

QY 193 TMDKNQKVLPAVAAGSGLLAHPNSYVQGPLSQLGFKKALSDVDTGKSLKGLKGPYLOVN 252

Db 184 EQSTLSKSAFAGISRAGFINNEDTFMGQFLIKMGIQPEVTKDKTTHVGERKGGPYIYN 243

QY 253 TETLSQVNPFRMFMNTN-KASSNEPSLKELEKDP-VWKKLNNAVKNQVRDILDRDLWARSR 310

Db 244 NEELANINPKVMILATDGTGDKNRTKF----IDPAVWKSLSKAVKDNKVYDVRNKLKSR 299

QY 311 GLISSEMAKELVELSKK 328

Db 300 GIIASESMAEDLEKIAEK 317

RESULT 10

Q7A6B6 PRELIMINARY; PRT; 319 AA.

ID Q7A6B6

AC Q7A6B6;

DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE SA0891 protein.

GN OrderedLocusNames=SA0891;

OS Staphylococcus aureus (strain N315).

OC Bacteria; Firmicutes; Bacillales; Staphylococcus.

OX NCBI_TaxID=158879;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21311952; PubMed=11418146; DOI=10.1016/S0140-6736(00)04403-2;

RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I., Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T., Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A., Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C., Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J., Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.

RT "Whole genome sequencing of methicillin-resistant Staphylococcus aureus."

RL Lancet 357:1225-1240(2001).

DR EMBL; AP003132; BAB42136.1; -.

DR GO; GO:0005381; F:iron ion transporter activity; IEA.

DR GO; GO:0006827; P:high affinity iron ion transport; IEA.

DR InterPro; IPR002491; Peripla_BP.

DR Pfam; PF01497; Peripla_BP_2; 1.

KW Complete proteome.

SQ SEQUENCE 319 AA; 35566 MW; DE6D7C7AB4831574 CRC64;

Query Match 42.9%; Score 724.5; DB 2; Length 319;

Best Local Similarity 46.9%; Pred. No. 6.5e-32;

Matches 149; Conservative 63; Mismatches 95; Indels 11; Gaps 4;

QY 13 VIGLLFVLIIATAACGNSSSSSKSSKSDGVEIKHEEGTTKVPKPKRVVVLVLEYSFVDAL 72

Db 9 VVFMILLVAVAGCGQKDT-----EKEETMTIKDELGTEKIKKPKRVVVLVLEYSFADYL 63

QY 73 VALDVKPVGIADDNKKRRIKPLRDKIGKYSVGTGRKQNPNEISIKLKPDLIIADNNRHK 132

Db 64 AALDMKPVGIADGSGTKNITKSVDRKIGAYESVGSRPQNMVEISKLPDLIIADVSRHK 123

QY 133 GIYKDLNKIAPTIELKSPDGYNENIDAKTISKALGKEEGKRLKEHDKKIEEYKKEI 192

Db 124 KIKSELSKIAPTIMLVSGTGDYNANIEAFKTVAKAVGKEGKELKHKDLAEIRKKI 183

QY 193 TMDKNQKVLPAVAAGSGLLAHPNSYVQGPLSQLGFKKALSDVDTGKSLKGLKGPYLOVN 252

Db 184 EQSTLSKSAFAGISRAGFINNEDTFMGQFLIKMGIQPEVTKDKTTHVGERKGGPYIYN 243

QY 253 TETLSQVNPFRMFMNTN-KASSNEPSLKELEKDP-VWKKLNNAVKNQVRDILDRDLWARSR 310

Db 244 NEELANINPKVMILATDGTGDKNRTKF----IDPAVWKSLSKAVKDNKVYDVRNKLKSR 299

QY 311 GLISSEMAKELVELSKK 328

Db 300 GIIASESMAEDLEKIAEK 317

RESULT 11

Q6GI45 PRELIMINARY; PRT; 319 AA.

ID Q6GI45

AC Q6GI45;

DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Transport system extracellular binding lipoprotein.

GN OrderedLocusNames=SA01011;

OS Staphylococcus aureus (strain MRS252).

OC Bacteria; Firmicutes; Bacillales; Staphylococcus.

OX NCBI_TaxID=282458;

RN [1]

RP SEQUENCE FROM N.A.

RX PubMed=15213324; DOI=10.1073/pnas.0402521101;

RA Holden M.T.G., Feil E.J., Lindsay J.A., Peacock S.J., Day N.P.J., Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,

```

RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,
RA Feltwell T., Lennard Z., Harris B., Hauser H., Holroyd S., Jagels K.,
RA James K.D., Lennard N., Line A., Mayes R., Hoyle S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitch E., Rutherford K.M., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,
RA Spratt B.G., Parkhill J.,
RT "Complete genomes of two clinical Staphylococcus aureus strains:
RT evidence for the rapid evolution of virulence and drug resistance."
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
DR EMBL; BX571856; CAG40015.1; -.
DR GO; GO:0005381; P:iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2; 1.
KW Complete proteome.
SQ SEQUENCE 319 AA; 35566 MW; 5885A83E3F502ABE CRC64;

Query Match 42.6%; Score 719.5; DB 2; Length 319;
Best Local Similarity 46.2%; Pred. No. 1.2e-31;
Matches 147; Conservative 65; Mismatches 95; Indels 11; Gaps 4;

QY 13 VIGLLFVLIIATAACGNNSSSSKSSKDGVEIKHEEGTTKVPKPKRVVVLVSFDAL 72
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 9 VVFMILVVAVAGCQKDT-----BEKTEMTIKDELCTEKIKKPKRIVLVLEYSFADYL 63
QY 73 VALDVKPGVIADDNKKRIIPKLRDKIGKYSVGTGRKQPNLEISKLKPDLLIADNNRHK 132
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 64 AALDKMPGVIADDGSKTKNITKSVRDKIGAYESVSRPQPNMEVLSKLKPDLLIADVSRHK 123
QY 133 GIYDKLNKIAPTIELKSPDGYNENIDAPKTSIKALGKEEGSKLEEDHKIKIEYKKEI 192
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 124 KIKSELSKIAPTILVSGTDYNNANIEAPTKAVAKVGKEGSKLEKHNKILABIRKKI 183
QY 193 TMDKNQKVLPAVAAGSLLAHPNSYVGQFLSQLGFEALSDDVTGKLSKYLKGPYLOWN 252
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 184 EQSTLSAPAGISAGFINNEDTFMQGFLIKMGIQPEVTKDHTAHERGKGPGYIYNL 243
QY 253 TETLSQVNPFRFIMTN-KASNEPSLKELEKDP-VWKKLNKAVNQKRVVDILDRDLWAKSR 310
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 244 NEELANINPKVMILATNGTKDKNTRKPF---IDPAVMKSLKAVKONKVVVDNRNKLQSR 299

QY 311 GLISSEENAKVELSKK 328
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 300 GIMASESMAEDLEKIAEK 317

RESULT 12
C03438 PRELIMINARY; PRT; 315 AA.
ID C03438 Q79ET3;
AC 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE YfmC protein.
GN Name=YfmC; OrderedLocusNames=BSU07520;
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]_TaxID=1423;
RP SEQUENCE FROM N.A.
RC STRAIN=166;
RX MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;
RA Kunat F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertsch M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Comercon I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Enlian K.-B., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,

```

```

RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M.,
RA Jones L.-M., Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M.,
RA Klein C., Kobayashi Y., Koetter P., Konigstein G., Krohn S.,
RA Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,
RA Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Maue C.,
RA Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S.,
RA Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,
RA Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,
RA Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G.,
RA Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B.,
RA Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeter R.,
RA Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Serror P.,
RA Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,
RA Takemaru K., Takeuchi M., Takakoshi A., Tanaka T., Terpstra P.,
RA Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F.,
RA Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H.,
RA Weitzenecker T., Winters P., Wipat A., Yamamoto H., Yamane K.,
RA Yasumoto K., Yata K., Yoshida K., Yoshikawa H.P., Zumstein E.,
RA Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis".
RL Nature 390:249-256(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AC327;
RX MEDLINE=97417488; PubMed=9272861; DOI=10.1016/S0378-1119(97)00130-3;
RA Yamamoto H., Uchiyama S., Nugroho F.A., Sekiguchi J.,
RT "Cloning and sequencing of a 35.7 kb in the 70 degree-73 degree region
RT of the Bacillus subtilis genome reveal genes for a new two-component
RT system, three spore germination proteins, an iron uptake system and a
RT general stress response protein."
RL Gene 194:191-199(1997).
DR EMBL; Z99108; CAB12561.1; -.
DR EMBL; D86417; BAA22317.1; -.
DR PIR; B69812; B69812.
DR GO; GO:0005381; P:iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2; 1.
KW Complete proteome.
SQ SEQUENCE 315 AA; 35020 MW; 2DA3BFFC49EFC39 CRC64;

Query Match 40.2%; Score 679.5; DB 2; Length 315;
Best Local Similarity 47.0%; Pred. No. 1.8e-29;
Matches 149; Conservative 50; Mismatches 103; Indels 15; Gaps 4;

QY 8 LKILSVIGLLEVLIIATAACGNNSSSSKSSKDGVEIKHEEGTTKVPKPKRVVVLVS 67
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 9 IAIMSVL-LLACLIVSGCSSQNNNGSKSKDSRVIHDEGKTTVSGTKRVVVLVS 67
QY 68 FVDALVALDVKPGVIADDNKKRIIPKLRDKIGKYSVGTGRKQPNLEISKLKPDLLIAD 127
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 68 FLDAVHNLGITPVGVIADDNKKDMIKLVGSSI-DYTSVGTSRSEPNLEIVISLKPDLIIAD 126
QY 128 NNRHKGIVKLNKIAPTIELKSPDGYNENIDAPKTSIKALGKEEGSKLEEDHKIKIEE 187
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 127 AERHKNYIKQLKTIAPTIELKSPDGYNENIDAPKTSIKALGKEEGSKLEEDHKIKIEE 186
QY 188 YKKEITMDKNQKVLPAVAAGSLLAHPNSYVGQFLSQLGFEALSDDVTGKLSKYLKGP 247
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 187 LKAEIPKDNENIVLGVARAOSFQLHTSSSYDGEIFKMLGFTTHAVKSD-----NA 236
QY 248 YLQWNTETLSQVNPFRFIMTN-KASNEPSLKELEKDP-VWKKLNKAVNQKRVVDILDRDLWA 307
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 237 YQEVSLQSLKIDPDILFISANEKGT---IVDEWKTNPWLKNLKAQKNGQVYDADRDTWT 293
QY 308 RSRGLISSEENAKVELIVE 324
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 294 RFRGKISSETSAKQVLK 310

RESULT 13
O07616

```


[illegible]

```

Db      200 SVLAALGL-----SVKPKIN---NAPWASINLQLLAINPQWLIVTHYR---EESIVKRW 247
QY      281 LEKDPVKKKLNKVNQRVDILDRDLWARSRLGISSEEMAKELVELSK 327
Db      248 WOODTLWNLEAQKQKQIAAVDSNAWARMRGIFAAERVGSDTVKIFK 294

RESULT 20
Q93F19 PRELIMINARY; PRT; 300 AA.
AC Q93F19,
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
FeCB.
DE GN Name=fecB;
OS Shigella flexneri 2a.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=42897;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=21437601; PubMed=11553538;
RX DOI=10.1128/IAI.69.10.6012-6021.2001;
RA Luck S.N., Turner S.A., Rajakumar K., Sakellaris H., Adler B.;
RT "Ferric dicitrate transport system (fec) of Shigella flexneri 2a
RT YSH6000 is encoded on a novel pathogenicity island carrying multiple
RT antibiotic resistance genes.";
RT Infect. Immun. 69:6012-6021 (2001).
RN [2]
RN SEQUENCE FROM N.A.
RA Luck S.N., Turner S.A., Rajakumar K.;
RA Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF326777; AAL08454.1; -.
DR GO; GO:0005381; F:iron ion transporter activity; IEA.
DR GO; GO:0008622; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP.2; 1.
DR SEQUENCE 300 AA; 33166 MW; 6E0B0CF5B9994AAD CRC64;

Query Match 27.4%; Score 463; DB 2; Length 300;
Best Local Similarity 35.5%; Pred. No. 1.1e-17;
Matches 100; Conservative 62; Mismatches 104; Indels 16; Gaps 4

QY 45 IKHEGTTKPKGKRVVVLEYSFVDALVALDVDPVGIADDNKKNRIIKPLRDKIGKYS 104
Db 24 VDEHGTTLEKTPQRIVVLELSFADALAADVSPIGIADDDNAKRILPEVRAHLKPWOS 83

QY 105 VGTQRQPNLEISLKPDLIIADNNRHGIIYKDINKTIAPTELKSFQDGYNNIDAFKTI 164
Db 84 VGTQAQPSLEIAALKPDLIIADSSRHAGVYIALQIAPVLVLLKSRNETVAENLQSAAI 143

QY 165 SKALQSEBEGKKLEHDKKLEEVKKEITMDKQKVLPAVAKSGLLAHPSNSYVQQLS 224
Db 144 GEMVQKGEQMARLEQHKERMAQWASQ--LPKGTRVAFGTSRSEQQFNLHTQETWGSVLA 201

QY 225 QLQGFKEALSDDTKGLSKYLKGPYL-QWNTETLSQVNPFRMFTMTNKAASNEPSLKELEK 283
Db 202 SLGL-----NVPMAAGSSMPSIGLEQLLAVNPALLVAHYR---EESIVKRWQ 248

QY 284 DPVKKKLNKVNQRVDILDRDLWARSRLGISSEEMAKELVEL 325
Db 249 DPLWMLTAQKQOAVSDSNWARMRGIFAAERIAADTVKI 290

```

Name=fecB; OrderedLocusNames=b290;
 Escherichia coli.
 Bacteria; Proteobacteria; Gammaproteobacter
 Enterobacteriaceae; Escherichia.
 NCBI_TaxID=562;
 [1]
 SEQUENCE FROM N.A.
 STRAIN=X12;
 MEDLINE=89213950; PubMed=2651410;
 Staudenmaier H., van Hove B., Yaraghi Z., Braun V.
 "Nucleotide sequences of the fecBCE genes and local
 proteins suggest a periplasmic-binding-protein-depend
 mechanism for iron(III) dicitrate in Escherichia coli."
 J. Bacteriol. 171:2626-2633(1989).
 [2]
 SEQUENCE FROM N.A.
 STRAIN=X12 / MG1655;
 MEDLINE=95334362; PubMed=7610040;
 Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
 Blattner F.R.;
 "Analysis of the Escherichia coli genome VI: DNA sequence of the
 region from 92.8 through 100 minutes."
 Nucleic Acids Res. 23:2105-2119(1995).
 -!- FUNCTION: Binds citrate-dependent iron(III); part of the binding-
 protein-dependent transport system for uptake of citrate-dependent
 iron(III).
 -!- SUBCELLULAR LOCATION: Periplasmic.
 -!- SIMILARITY: Belongs to the bacterial extracellular solute-binding
 protein family 8.
 -!- SIMILARITY: Contains 1 iron siderophore/cobalamin periplasmic-
 binding domain.

Query Match	27.4%;	Score 463;	DB 2;	Length 300;
Best Local Similarity	35.5%;	Pred. No. 1.1e-17;		
Matches 100;	Conservative 62;	Mismatches 104;	Indels 16;	Gaps 4

QY	45	IKHEGGTTKPKPKRVVVLVEYSFDALVALDVKPVGIADDNKNRIIKPLRDKICKYTS	104
DB	24	VQDEHGTTLEKTPQRIIVVLELSPADALAAVDVSPIGIADNDNAKRILPEVRAHLKPMWS	83
QY	105	VGTRKQPNLEISKLKPDIIADNNRHKGIIYKDLINKAPIETLKSFGDYNENIDAFKTI	164
DB	84	VGTRAQPSLEIAIALKPDIIADSSRHAGVYIALQIAPVLLLLKSRNETVAENLQSAAI	143
QY	165	SKALCKEGBEGKRLSEHDKKIEEYKKEITWDDKNQKVLPAVAKSGULLAHPSSNYVGQFLS	224
DB	144	GEWQCKGEMQARLEHSEKHAQWASQ--LPKGRVAFGTSRQQQFNLIHTQETWGSVLA	201
QY	225	QLGFKREALSDVTKGLSKYLGXPYL-QMNTETLSQVNPERFMFTMTKASNEPSLKELK	283
DB	202	SLGL-----NVPAAMAGSMPSIGLEQLLAVNPAMLLVAHYR---EESIVKRWQQ	248
QY	284	DPVKKKLNAVKNQVDIILDRDLMAWSRGLSSISEEMAKELVEL	325
DB	249	DPLQMLTAQAKQKQVQVSDSNTWARMRGIFAERIAADTVKI	290

RESULT 21
FECS_ECOLI
ID FECS_ECOLI
AC P15028; P76816;
DT 01-APR-1990 (Rel. 14, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Iron(III) dicitrate-binding periplasmic protein precursor.

Db 144 GEMVCKGKREMOARLEQHKERMAQWASQ--LPKGTTRVAFCTSRQEQFNHLHQTBTWTGSLVA 201
Qy 225 QLGPKREALSDVTGKGLSKYLGKPYL-QMNTETLSQVNPFRMFTMTKASSNEPSSLKELEK 283
Db 202 SLGL-----NVPAMAGASPSGLGLEQLLAVNPAWLLVAHYR---EESIVRRWQ 248
Qy 284 DPVWKKLNAVKNQVRDILDRDLWARSRLGSLISEEMAKELVEL 325
Db 249 DPLWQMLTAQKQVAVSDNTWARMRGIFAAERIAADTVKI 290
RESULT 22
Q88916 PRELIMINARY; PRT; 306 AA.
AC Q88916;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Iron(III) dicitrate transport system, periplasmic iron-binding protein
DE PecB.
GN Name=fecB; OrderedLocusNames=PSPT00763;
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DC3000;
RX MEDLINE=22834015; PubMed=12928499; DOI=10.1073/pnas.1731982100;
RA Buell C.R., Joarder V., Lindeberg M., Selenut J., Paulsen I.T.,
RA Gwin M.L., Dodson R.J., DeBoy R.T., Durkin A.S., Kolonay J.F.,
RA Madupu R., Dougherty S.C., Brinkac L.M., Beanan M.J., Haft D.H.,
RA Nelson W.C., Davidson T.M., Zafar N., Zhou L., Liu J., Yuan Q.,
RA Khouri H.M., Fedorova N.B., Tran B., Russell D., Berry K.J.,
RA Utterback T.R., Van Aken S.E., Feldblyum T.V., D'Ascenzo M.,
RA Deng W.-L., Ramos A.R., Alfano J.R., Cartinhour S., Chatterjee A.K.,
RA Delaney T.P., Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X.,
RA Bender C.L., White O., Fraser C.M., Collier A.;
RT "The complete genome sequence of the Arabidopsis and tomato pathogen
RT Pseudomonas syringae pv. tomato DC3000.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186 (2003).
RL EMBL; AE016858; AA054305.1; -;
DR TIGR; PSPT00763; -;
DR GO; GO:0005381; F:iron ion transporter activity; IEA.
DR GO; GO:0005827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2; 1.
KW Complete proteome.
SQ SEQUENCE 306 AA; 32872 MW; 371BA711BC7A6108 CRC64;
Query Match 26.9%; Score 454.5; DB 2; Length 306;
Best Local Similarity 34.6%; Pred. No. 3.2e-17;
Matches 112; Conservative 60; Mismatches 111; Indels 41; Gaps 7;
Qy 11 LSVIGLLFVLIAATAACGNNSSSSKSSKSGDVGVEIKHEEGTKV--PKHPKRVVLEYSF 68
Db 9 LLACGLL-TLLSAA-----QAAPDIDGQHKVLPDTPKRVVLEFSF 52
Qy 69 VDVALVDKVPVGIADNDKNRIIPKLDKIGKYSVGTGRKQPNLEESKLKPDLLIADN 128
Db 53 LDGLASGVTPVGAADGDASRVLPKRVKAVGEMQSVGLRSQPNIEVIARLKPDLIIADL 112
Qy 129 NRHKGKYLKNTAPTIKELSGDYNENIDAFKTSKALGKEEGKKEEHDHKKIEY 188
Db 113 GRHQALYNDLASLAPTLMLPSRGEDYQGLSKSAGLIGMALGKGPENQARIAENRHLKTV 172
Qy 189 KKEITMDKNQKVLPAVAAGSGLLAHPSNSVVGQFSLQGLKALSDDVTGKLSKYLKGPY 248
Db 173 AEQIPADSN--VLFGVAREDSFVHGPHSYAGSVLQAIG-----LQVPE 214
Qy 249 LQMTETLSQVNPFRMFTMTK-----ASSNEPSSLKEL-EKDPVWKKLNAVKNQVRDILDR 303
Db 215 VRNNAAPTEFVSLQQLALDPNLLVGHVRRPSIVDTWSKQPLWQVLGAVRNQVAEVDG 274

Qy 304 DLWARSRLGSLISEEMAKELVELSK 327
Db 275 DSWARNRGIMASEQIADDALAVLK 298
RESULT 23
Q85F62 PRELIMINARY; PRT; 325 AA.
AC Q85F62;
DT 25-OCT-2004 (TRENBLrel. 28, Created)
DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)
DE YfiY (Periplasmic binding protein).
GN Name=yfiY; ORFNames=BL02129, BL103475;
OS Bacillus licheniformis DSM 13.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=279010;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 13;
RX PubMed=15383718;
RA Veith B., Herzberg C., Steckel S., Feesche J., Maurer K.H.,
RA Ehrenreich P., Baumer S., Henne A., Liebegang H., Merkl R.,
RA Ehrenreich A., Gottschalk G.;
RT "The Complete Genome Sequence of Bacillus licheniformis DSM13, an
RT Organism with Great Industrial Potential.";
RL J. Mol. Microbiol. Biotechnol. 7:204-211(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 14580;
RA Rey M.W., Ramaiya P., Nelson B.A., Brody-Karpin S.D., Zaretsky E.J.,
RA Tang M., de Leon A.L., Xiang H., Gusti V., Clausen I.G., Olsen P.B.,
RA Rasmussen M.D., Andersen J.T., Jorgensen P.L., Larsen T.S.,
RA Sorokin A., Bolotin A., Lapidus A., Galleron N., Ehrlich S.D.,
RA Berka R.M.;
RT "Complete genome sequence of the industrial bacterium Bacillus
RT licheniformis and comparisons with closely related Bacillus species.";
RL Genome Biol. 5:R77-R77(2004).
DR EMBL; AE017333; AAU42302.1; -;
DR EMBL; CP000002; AAU4933.1; -;
SQ SEQUENCE 325 AA; 36700 MW; 949283E3C6C24775 CRC64;
Query Match 24.6%; Score 415.5; DB 2; Length 325;
Best Local Similarity 33.2%; Pred. No. 4.6e-15;
Matches 108; Conservative 63; Mismatches 117; Indels 37; Gaps 10;
Qy 8 LKLSVIGLLFVLIAAT--AACGNNSSSSKSSKSGDVGVEIKHEEGT-TKVPKHPKRVVVL 64
Db 1 MKRWSIVGFIALLAISILAAACGGKDDSAKGGKNDTITVKHMGTEQVDPADPKRVVVL 60
Qy 65 EYSFVDALVALDVKPVGIADNDKNRIIPKLDKIGKYSVGTGRKQPNLEESKLKPDLLI 124
Db 61 TNEGTEALLELGKVPVAVKSVGTGDPWYKHQIADMKGVKQIGLEGEPNIEAIAELKPDLLI 120
Qy 125 IADNNRHGIYKDLNKIAPTIELKSGFDGYNENIDAFKTSKALGKEEGKKEEHDHKK 184
Db 121 IGKWRHEKLYDKLEKIAPTVEFTLRGEWKDN---FSLYAKAVNKEEGKKEEHDHKK 177
Qy 185 IBEYKKEITMDKNQK--VLPAVAAGSGLLAHPSNSVVGQFSLQGLK-----EAL---SD 234
Db 178 VSELSDKLGDQKKRVSVIRFVAGQSRI--YYKDSFPFGIILDLQGLGFKYPEKQELFEKQD 235
Qy 235 DVTGKLSKYLKGPYLOMNTETLSQVNPFRMFTMTKASSNEPSSLKELE-----KDPVW 287
Db 236 D-----KPVFTENKESIPDDMGDLVFTFYTKPAB---SKKEAETQNDWTSDDLW 283
Qy 288 KKLNAVKNQVRDILDRDLWARSRL 312
Db 284 KKLNAVKNQVNAHEVDALNTTAGGV 308
RESULT 24

```

O87491
ID O87491 PRELIMINARY; PRT; 330 AA.
AC O87491;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 23-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Lipoprotein SirA.
GN Names:SirA;
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=ISP3;
RX MEDLINE=99175440; PubMed=10049373;
RA Heinrichs J.H., Gatlin L.E., Kunsch C., Choi G.H., Hanson M.S.;
RT "Identification and characterization of SirA, an iron-regulated
protein from Staphylococcus aureus.";
RL J. Bacteriol. 181:1436-1443 (1999).
RN [2]
SEQUENCE FROM N.A.
RP PubMed=14688077; DOI=10.1128/TAI.72.1.29-37.2004;
RX Dale S.E., Doherty-Kirby A., Lajoie G., Heinrichs D.E.;
RT "Role of siderophore biosynthesis in virulence of Staphylococcus
aureus: identification and characterization of genes involved in
production of a siderophore.";
RL Infect. Immun. 72:29-37 (2004).
DR EMBL; AF079518; AAC62496.1; -
DR EMBL; AY251022; AAP82062.1; -
DR PIR; G89771; G89771.
DR GO; GO:0005381; F:iron ion transporter activity; IEA.
DR GO; GO:0006837; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_Bp.
DR Pfam; PF01497; Peripla_BP_2; 1.
KW Lipoprotein.
SQ SEQUENCE 330 AA; 36744 MW; 63710BA22B208F61 CRC64;

Query Match 24.0%; Score 405; DB 2; Length 330;
Best Local Similarity 32.8%; Pred. No. 1.7e-14;
Matches 113; Conservative 65; Mismatches 132; Indels 34; Gaps 12

Qy 8 LKLSVIGLLPVLVIATAACGNSSNSKSGVETKHEGTTKVPKPKRVVVLEYS 67
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 5 IKML-VVTLAFLLV-LAGCSGNKQSSDNKDQKETTISKHAGTTTEIKGPKRVVTLYG 62
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 68 FVDALVALDVKPGVIADNKNRIKPRDKIGKTVSGTRKQPNLEISKLPDLIAD 127
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 63 ATDVAVSLGVKPGVAVESWTKQKPEYIKNDLKDTKIVGQEPAPNLEISKLPDLIVAS 122
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 128 NNRHKGIVKIDINKTAPTIETLKSFDGYNENIDAFKTIKALGKEEGKKRLEHDKTIE 187
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 123 KVRNEKVDYQLSKIAPT-V---STDTVFKFK-DTTKLMGKALGKEAEADLLKKYDDKVA 178
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 188 YKKEITWDKQKVLPAVAASKGLLAHPSN-----SYVQFSLQGFKEALSDVTKGL 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 179 FQK---DAKAKYKDAWFLKASVNVNFRADHTRIYAGGYAGETLNDLGPKR-----NKDL 228
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 241 SKYLKG--PYLQMT-ETLSQVNPFRMIMTNKASNEPSL-----KELEKDPVVKLLNA 292
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 229 QKQVDNGKDIILQTSKESIPLMNADHIFVVKSDPNKDAALVKYKTESEWTSSEKWNKDLA 288
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 293 VKNQRV-DILDRDLWARSGRLISSSEMAKELVELSKDSKDNK 335
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 289 VKNQVSDDLDEITWNLAGGYKSSLKGLIDDLVE--KLNIEKQSK 330
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 25
Q7A1Z7 PRELIMINARY; PRT; 330 AA.
ID Q7A1Z7
AC Q7A1Z7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

```

[illegible]


```

QY 132 KGIYKDLNKIAPTIELKSPDGDYNNIDAFKTIISKALKEEGKRLBEHDKKIEEYKKE 191
Db 132 EKVEQLKAIAPTVPSETLGEWKDN---PKFYAKALNKEKGQKVLAAAYDKRMKDLKAK 188
QY 192 ITMDKNQKVLPAVAAKSGLLAHPNSVYGQFSLQSGFK----EALSDVDTKGLSKYLKGP 247
Db 189 LGDKVQNEISWVRPMPGVDVRIYHGTFTSGVILKELGFKRPGDQNKDPAERNVSK----- 243
QY 248 YLQMTETLSQVNPFRMFMINTKASSNEPSLKELEK---DPVWKKLNAVKNQVLDLRLD 303
Db 244 -----ERISAMGDGVLFYFTFD-KGNEKKGSELEKEYINDPLFKNLNAVKNQKAYKVDD 296
QY 304 DLWARSGLISS-----EEMAKELVE 324
Db 297 VIVNTAGGVMAANLLDDDDIEKRFVK 321

RESULT 31
Q6HNJ9 PRELIMINARY; PRT; 321 AA.
AC Q6HNJ9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Iron(III) dicitrate ABC transporter, periplasmic protein.
GN Name=fecB; OrderedLocusNames=B79727.0525;
OS Bacillus thuringiensis (subsp. konkukian).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=180856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97-27;
RA Brettin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Tice H.;
RT "Complete genome sequence of Bacillus thuringiensis 97-27.";
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB017355; AAF62409.1; --
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2; 1.
KW Complete proteome.
SQ SEQUENCE 321 AA; 36373 MW; 47P614F471C1CF3 CRC64;

Query Match 23.9%; Score 404; DB 2; Length 321;
Best Local Similarity 31.4%; Pred. No. 1.9e-14;
Matches 101; Conservative 73; Mismatches 120; Indels 28; Gaps 7;

QY 16 LLFVLIIATAACGNSSSSSKE-SSKDGVEIKHEGTTKVPKPKRVVLEYSFVDALVA 74
Db 15 LLAFSLLSACGKNTKESKEDTKKEMIPVDHANGKTEVPANPKRVVILNEGTEALLE 74
QY 75 LDVPGVGIADNKNRIIPLRDKIGKYSVGTGRKQPNLEESIKLPDLIIADNNRHKGI 134
Db 75 LGVPGVAGKSVTGDWPYPHIKDKMKVKVGDGVQNVETIASLKPDLIIIGNKVRHEKV 134
QY 135 YKDLNKIAPTIELKSPDGDYNNIDAFKTIISKALKEEGKRLBEHDKKIEEYKKEITM 194
Db 135 YEQLKAIAPTVPSETLGEWKDN---PKFYAKALNKEKGQKVLAAAYDKRMKDLKAKLGD 191
QY 195 DKNQKVLPAVAAKSGLLAHPNSVYGQFSLQSGFK----EALSDVDTKGLSKYLKGPYLO 250
Db 192 KVNQEISWVRPMPGVDVRIYHGTFTSGVILKELGFKRPGDQNKDPAERNVSK----- 243
QY 251 MNTETLSQVNPFRMFMINTKASSNEPSLKELEK---DPVWKKLNAVKNQVLDLRLD 306
Db 244 ---ERISAMGDGVLFYFTFD-KGNEKKGSELEKEYINDPLFKNLNAVKNQKAYKVDDVIV 299
QY 307 ARSRGLISS-----EEMAKELVE 324
Db 300 NTAGGVMAANLLDDDDIEKRFVK 321

RESULT 33
Q73DM9 PRELIMINARY; PRT; 321 AA.
AC Q73DM9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Iron compound ABC transporter, iron compound-binding protein.
GN OrderedLocusNames=BCE0683;
OS Bacillus cereus (strain ATCC 10987).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=22523;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14960714; DOI=10.1093/nar/gkh258;
RA Rasko D.A., Ravel J., Oekstad O.A., Helgason E., Cer R.Z., Jiang L.,
RA Shores K.A., Fouts D.E., Tourasse N.J., Angiuoli S.V., Kolonay J.F.,
RA Nelson W.C., Kolstoe A.-B., Fraser C.M., Read T.D.;
RT "The genome sequence of Bacillus cereus ATCC 10987 reveals metabolic

```

```

RESULT 32
Q63G31 PRELIMINARY; PRT; 321 AA.
AC Q63G31;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Iron(III) dicitrate ABC transporter, periplasmic protein.
GN Name=fecB; ORFNames=BTZK0525;
OS Bacillus cereus ZK.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=288681;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZK;
RA Brettin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Tice H.;
RT "Complete genome sequence of Bacillus cereus ZK.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CP000001; AAU19716.1; --
DR EMBL; CP000001; AAU19716.1; --
SQ SEQUENCE 321 AA; 36315 MW; 0A1B4742B1AB7945 CRC64;

Query Match 23.6%; Score 399; DB 2; Length 321;
Best Local Similarity 31.4%; Pred. No. 3.6e-14;
Matches 101; Conservative 71; Mismatches 122; Indels 28; Gaps 7;

QY 16 LLFVLIIATAACGNSSSSSKE-SSKDGVEIKHEGTTKVPKPKRVVLEYSFVDALVA 74
Db 15 LLAFSLLSACGKNTKESKEDTKKEMIPVDHANGKTEVPANPKRVVILNEGTEALLE 74
QY 75 LDVPGVGIADNKNRIIPLRDKIGKYSVGTGRKQPNLEESIKLPDLIIADNNRHKGI 134
Db 75 LGVPGVAGKSVTGDWPYPHIKDKMKVKVGDGVQNVETIASLKPDLIIIGNKVRHEKV 134
QY 135 YKDLNKIAPTIELKSPDGDYNNIDAFKTIISKALKEEGKRLBEHDKKIEEYKKEITM 194
Db 135 YEQLKAIAPTVPSETLGEWKDN---PKFYAKALNKEKGQKVLAAAYDKRMKDLKAKLGD 191
QY 195 DKNQKVLPAVAAKSGLLAHPNSVYGQFSLQSGFK----EALSDVDTKGLSKYLKGPYLO 250
Db 192 KVNQEISWVRPMPGVDVRIYHGTFTSGVILKELGFKRPGDQNKDPAERNVSK----- 243
QY 251 MNTETLSQVNPFRMFMINTKASSNEPSLKELEK---DPVWKKLNAVKNQVLDLRLD 306
Db 244 ---ERISAMGDGVLFYFTFD-KGNEKKGSELEKEYINDPLFKNLNAVKNQKAYKVDDVIV 299
QY 307 ARSRGLISS-----EEMAKELVE 324
Db 300 NTAGGVMAANLLDDDDIEKRFVK 321

RESULT 33
Q73DM9 PRELIMINARY; PRT; 321 AA.
AC Q73DM9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Iron compound ABC transporter, iron compound-binding protein.
GN OrderedLocusNames=BCE0683;
OS Bacillus cereus (strain ATCC 10987).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=22523;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14960714; DOI=10.1093/nar/gkh258;
RA Rasko D.A., Ravel J., Oekstad O.A., Helgason E., Cer R.Z., Jiang L.,
RA Shores K.A., Fouts D.E., Tourasse N.J., Angiuoli S.V., Kolonay J.F.,
RA Nelson W.C., Kolstoe A.-B., Fraser C.M., Read T.D.;
RT "The genome sequence of Bacillus cereus ATCC 10987 reveals metabolic

```

RT adaptations and a large plasmid related to *Bacillus anthracis* pXOI.1";
RL Nucleic Acids Res. 32:977-988 (2004).
DR EMBL; AB017266; AAS39616.1; -.
DR TIGR; BC80683; -.
DR GO; GO:0005381; F:iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2; 1.
DR Complete proteome.
KW Complete proteome.
SQ SEQUENCE 321 AA; 36322 MW; ACDEA9056A566FA5 CRC64;

Query Match 23.5%; Score 397; DB 2; Length 321;
Best Local Similarity 31.0%; Pred. No. 4.6e-14;
Matches 100; Conservative 77; Mismatches 116; Indels 30; Gaps 8;

QY 16 LLFVLIIATAAGNNSSSSSKE--SSKDGVEIKHEGTTKVPKHPKRVVVLVEYSFWDALVA 74
DB 15 LLAFSLLLSACGSNTYKEDTKKMITVEHANGKTEVPANPKRVVILTNEGTEALLE 74
QY 75 LDVKPVGAIADNDKKNRIKPLRDKIGKYTSVGRKQPNLEEISKLKPLDIADNNRHGI 134
DB 75 LGVKPVGAVKSWTGDWPVPHIKDKMKDVKVVDGSGQVNVETIASLQPDLLIIGNKMRHEKV 134
QY 135 YKOLNKIAPTIELKSPFGDYNENIDAKFTISKALGKEEGKRLKEEHDKKIEEYKKEITM 194
DB 135 YEQLKAIAPTVESETLRLGEWKDN---PKFYAKALNKEKEGQKVADLYEKSMDLKAKLGD 191
QY 195 DKQKQVLPAVAKSGLLAHPSSYVVGQFLSLQGFK---EALSDDVTKGLSKYLKGPVLQ 250
DB 192 KVNQIEISNVRFPMDGVRIYHGDITFSGVILKELGFRPGDQNKQDPAERNVSK----- 243
QY 251 MNTETLSQVNPERRFIMT-NKASNEBPSLKELEK-----DPVWKKLNAVONORVDILDRDL 305
DB 244 ---ERISAMGDGVLFFYTFDKGNENKGS--ELEKEYINDPLFKNLNAVKNKGKAYKVDDVI 298
QY 306 WARSRLGISS-----EEMAKELVE 324
DB 299 WNTAGGVMAANLLDDIEKGRFVK 321

RESULT 34
Q81109 PRELIMINARY; PRT; 322 AA.
ID AC Q81109;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Iron(III) dicitrate-binding protein.
GN OrderedLocusNames=BC0616;
OS *Bacillus cereus* (strain ATCC 14579 / DSM 31).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=226900;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22608415; PubMed=12721630; DOI=10.1038/nature01582;
RA Ivanova N., Sorokin A., Anderson I., Galleron N., Candelson B.,
RA Kapatal V., Bhattacharyya A., Reznik G., Mikhailova N., Lapidus A.,
RA Chu L., Mazur M., Goltzman E., Larsen N., D'Souza M., Walunas T.,
RA Grechkin Y., Pusch G., Haseelkorn R., Fomstein M., Ehrlich S.D.,
RA Overbeek R., Kyripides N.C.;
RT "Genome sequence of *Bacillus cereus* and comparative analysis with
RT *Bacillus anthracis*.";
RL Nature 423:87-91(2003).
DR EMBL; AB017000; AAP07634.1; -.
DR GO; GO:0005381; F:iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2; 1.
DR Complete proteome.
KW Complete proteome.
SQ SEQUENCE 322 AA; 36417 MW; 760088C3D5D6E9FE CRC64;

Query Match 23.5%; Score 396.5; DB 2; Length 322;
Best Local Similarity 31.0%; Pred. No. 4.9e-14;

```
Db 115 PDLIITASPRGAIRKNELEQIAPTVMFDPSTSNNDHFAEMTETFKQIAKAVGKEBEGKKV 174
Qy 178 LEEHKKKEEYKKEI-----TMDKQKVLPAVAKS-----GLLAHNSNSVGVFLSQIGFK 229
Db 175 LADMKAFAFAKAKIEKADLKDKNIAQAQFAKQVPTFRILT--DNSLALQVTKKLG-- 230
Qy 230 EALSDDVTKGLSKYLKGPVLOMNTTSLQVNPFR-MFMTNKAASSNEPSLKELEKDPVWK 288
Db 231 --LTTTFEAKSE--PDGFKQTTVESLQSVQDSNFYIIVADEDNIFDTQLK---GNPAWE 283
Qy 289 KLNNAVKNQKVDILDRDLWARSRLISSEEMAKELVEL 325
Db 284 ELKPKKENKMYKLKGDWTWIFG-GPESATSLATQVADV 319

RESULT 36
Q81DY8 PRELIMINARY; PRT; 320 AA.
AC Q81DY8;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-WAR-2004 (TReMBLrel. 26, Last annotation update)
DE Iron(III) dicitrate-binding protein.
GN OrderedLocusNames=BC2208;
OS Bacillus cereus (strain ATCC 14579 / DSM 31).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=226900;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22608415; PubMed=1271630; DOI=10.1038/nature01582;
RA Ivanova N., Sorokin A., Anderson I., Galleron N., Candelon B.,
RA Kapral V., Bhattacharyya A., Reznik G., Mikhailova N., Lapidus A.,
RA Chu L., Mazur M., Gotsman E., Larsen N., D'Souza M., Walunas T.,
RA Grechkin Y., Pusch G., Haselkorn R., Fonstern M., Ehrlich S.D.,
RA Overbeek R., Kyrpides N.C.;
RT "Genome sequence of Bacillus cereus and comparative analysis with
RT Bacillus anthracis."
RL Nature 423:87-91(2003).
DR EMBL; AE017005; AAP09174.1; -.
DR GO; GO:0005381; F:iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
KW Complete proteome.
SQ SEQUENCE 320 AA; 36173 MW; 95ED63B62B07323F CRC64;

Query Match 23.1%; Score 391; DB 2; Length 320;
Best Local Similarity 30.2%; Pred. No. 9.7e-14;
Matches 98; Conservative 64; Mismatches 138; Indels 24; Gaps 5;

Qy 9 KILSVIGLLFVLIAATACG-NSSSSNSKSSKDGVEIKHEGTTKVPKPKRVVVLVS 67
Db 5 KKLIMFCIMLMVVIAGCSKEEKENNISAKADSVTVRHAMGETTVNGTPKRVVVLVTE 64
Qy 68 FVDALVALDVKPGIADNDKNRIIKPLDKIKGYTSVGRKPNLEISKLKPDILIAAD 127
Db 65 GAELALAVGTVPTGTTKPRAGDEWYHLAKELKNTVEVGTGRDINLEAVMKLPDILIGN 124
Qy 128 NNRHGIYKDLNKIAPTIELKSPDGYNNENIDAFKTSKALGKEBKGKRLBEHDKIEE 187
Db 125 KMRHEKIYDQLEIAPTVAETLRGDWKEN---FTLYTRAVNKEBKGKALDDYKKRIEG 181
Qy 188 YKEITMDKNQKVLPAVAKSGLLAHPNSVYQFLSQIGFK---EALSDDVTKGLSKY 243
Db 182 MKELGDKLNSKYSIRFVPGDVRIYQKNSFSGVGLNDIGFKRPPIQDKDDFAIKGITK- 240
Qy 244 LKGPYLOMNTTSLQVNPFRMFTM-----NKASSNEPSLKELEKDPVWKLNNAVKNQV 298
Db 241 -----EQIPNDGDYLFYFTSDKQADKNNEGNSLAKWETDPLFKQLQASKNKV 290
Qy 299 DILDRDLWARSRLISSEEMAKEL 322
Db 291 FQVDEVINWTAGGIIKAANLMLDDI 314
```

```
RESULT 37
Q81L65 PRELIMINARY; PRT; 324 AA.
AC Q81L65; Q6HSI7; Q6KLW1;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Iron compound ABC transporter, iron compound-binding protein.
GN OrderedLocusNames=BA4766; BAS4424; GBAA4766;
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=11392;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Ames / isolate Porton;
RX MEDLINE=22608414; PubMed=12721629; DOI=10.1038/nature01586;
RA Read T.D., Peterson S.N., Tourasse N.J., Baillie L.W., Paulsen I.T.,
RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,
RA Holtzapple E.K., Okstad O.A., Helgason E., Rilstone J., Wu M.,
RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.L.,
RA DeBoy R.T., Beanan M.J., Daugherty S.C., Durkin A.S., Haft D.H.,
RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,
RA Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,
RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,
RA Hazen A., Cline R.T., Redmond C., Thwaite J.E., White O.,
RA Salzberg S.L., Thomson B., Friedlander A.M., Koehler T.M.,
RA Hanna P.C., Kolstoe A.-B., Fraser C.M.;
RT "The genome sequence of Bacillus anthracis Ames and comparison to
RT closely related bacteria."
RL Nature 423:81-86(2003).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Ames / isolate 0581;
RA Ravel J., Rasko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,
RA Wilson M., Stanley S., Decker S., Read T.D., Salzberg S.L.,
RA Fraser C.M.;
RT "Bacillus anthracis comparative genomics."
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=Sterner;
RA Hittcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Tice H.;
RA Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; AE017038; AAP8458.1; -.
DR EMBL; AE017334; AAT33889.1; -.
DR EMBL; AE017225; AAT56722.1; -.
DR TIGR; BA4766; -.
DR TIGR; GBAA4766; -.
DR GO; GO:0005381; F:iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR008091; Ferrichrome_bind.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2; 1.
DR PRINTS; PRO1715; FERIBNDNGPP.
KW Complete proteome.
SQ SEQUENCE 324 AA; 36074 MW; 0EAD6197AB43B3A6 CRC64;

Query Match 23.1%; Score 391; DB 2; Length 324;
Best Local Similarity 34.4%; Pred. No. 9.8e-14;
Matches 116; Conservative 59; Mismatches 122; Indels 40; Gaps 13;

Qy 9 KILSVIGLLFVLIAATACGNSSSNSKSSKDXD---GVEIKHEGTTKVPKPKRVVVL 64
Db 3 KILSI----FIVVFLFVAVGCGQCGQCKEKTADNKNQAIITIKHAEGETKLDKPKRVVVL 59
Qy 65 EYSFVDALVALDVKPGIADNDKNRII---KPLDKIKGYTSVGRKPNLEISKLK 120
Db 59 EWYSEDLLALGVQVPVGMADIKNYNKWNVTKTPSKDVV----DVGTRQPNLEISRLK 114
Qy 121 PDLIITADNNRHKGIYKDLNKIAPTIEL---KSGFDGYNNENIDAFKTSKALGKEBKGK 177
```


Search completed: August 25, 2005, 09:23:53
Job time : 71 secs

This page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 09:19:06 ; Search time 20 Seconds
(without alignments)
1611.631 Million cell updates/sec

Title: US-10-724-972A-6352
Perfect score: 1690
Sequence: 1 GVSVRGLKTIISVIGLFFVL.....EEMAKELVELSKDSKONK 335

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 10%
                  Listing first 45

```

```
Database :      PIR_79:*
1:  _pir1:
2:  _pir2:
3:  _pir3:
4:  _pir4:
```

SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	1232	42.9	327	2	D90013	hypothetical prote
2	724.5	42.9	319	2	E89872	hypothetical prote
3	679.5	40.2	315	1	E69812	ferrichrome ABC tr
4	611	36.2	348	1	F69831	iron(III) dicitrat
5	460	27.2	302	1	QRECD3	citrate-dependent
6	405	24.0	330	2	E69771	lipoprotein (impor
7	373	22.1	325	1	C68905	iron(III) dicitrat
8	350	20.7	341	2	E83779	iron(III) dicitra
9	332.5	19.7	317	1	E69763	probable ferrichro
10	332.5	19.7	334	2	G84123	iron(III) dicitra
11	327	19.3	328	2	AH2128	iron(III) dicitrat
12	310	18.3	325	2	AG2082	iron(III) dicitrat
13	295.5	17.5	330	2	AG2129	iron(III) dicitrat
14	282.5	17.3	330	2	S7458	iron(III) dicitrat
15	288.5	17.1	319	2	AC2421	iron(III) dicitrat
16	288	17.0	326	2	AF2082	iron(III) dicitrat
17	285	16.9	300	2	AC1855	periplasmic iron-c
18	285	16.9	342	2	A89846	hypothetical prote
19	284	16.8	331	2	AD2134	iron(III) dicitrat
20	279.5	16.5	330	2	AD2082	iron(III) dicitrat
21	278	16.4	361	2	AE2074	ferrichrome bindin
22	277.5	16.4	319	2	AD2085	iron(III) dicitrat
23	277.5	16.4	357	2	AD2079	iron(III) dicitrat
24	276.5	16.4	317	2	I39842	iron-uptake system
25	275	16.3	315	2	S32930	ferrichrome ABC tr
26	274	16.2	343	2	AG2075	iron(III) dicitrat
27	274	16.2	350	2	T36412	probable iron-side
28	273	16.2	333	2	AB2130	iron(III) dicitrat
29	267.5	15.8	331	2	AI1943	iron(III) dicitrat

ALIGNMENTS

RESULT 1

D90013
hypothetical protein SA1979 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: D90013
R:A; Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, M.; Kuroda, M.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Seki, C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: D90013
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-327 <UNP>
A:Cross-references: KUNITROT:Q99589; GB:BA000018; PID:gl3701977; PIDN:BA043269.
A:Experimental source: strain N315
A:Genetics:
A:Gene: SA1979
A:Superfamily: ferrichrome-iron transport protein fecB

Query Match	72.9%;	Score 1232;	DB 2;	Length 327;
Best Local Similarity	71.9%;	Pred. No. 2.6e-62;		
Matches 235;	Conservative 42;	Mismatches 50;	Indels 0;	Gaps 0;
Qy	5	VRGLKILSVIGLLPVLITATACAGNNSNSSKSSKGVIEIKHBEGTTKVPKHPKRVVVL	64	
Db	1	MRGLKTSILGLIVALLVACAGNTDINSKESSTKDTISVDENGTVKVPDKAKRIVVL	60	
Qy	65	EYSFVDALVALDVKPVGIIADDNKKNRILIKPLURDKIGKTVSVGTRKQPNLEBISKLPDLIL	124	
Db	61	EYSFADALALDVKPVGIIADGKKRIIKPVREKIGDVTYSVGTRKQPNLEBISKLPDLIL	120	
Qy	125	IADNRHKGIIYKDLANKIAPTIELTKSFGDGYNNENIDAFKTIISKALGKEBEGKKLEBHHKK	184	
Db	121	IADSRHKGINKELNKIAPTILSLKSGFDGYQNINSFKTIIAKALNKEKEGEKRLAEHDKL	180	
Qy	185	IEEYKKEITMDKNOKVLPAVAAKSGLLAHPSNYSVVGQFLSOLGPKFKAALSDDDVTYKGLSKYL	244	
Db	181	INKYKDEITKFORNOKVLPAVVAKAGLLAHPNYSVVGQFLNELGPKNAALSDDDVTYKGLSKYL	240	
Qy	245	KGPLYQMNTETLSQVNPERPIMPITNKASNSBPISKLEKDPVWKKKLVATQNRQVDIRD	304	
Db	241	KGPLYQLDTEHLADLNPERMIIMTDHAKDSAEFKKLQEDATWKKLVATQNRVDIRD	300	
Qy	305	LWARSRLISSEEMAKELVELSKDOK	331	
Db	301	VWARSRLISSEEMAKELVELSKKEQ	327	

RESULT 2

C;Superfamily: ferrichrome-iron transport protein fecB

Query Match 18.3%; Score 310; DB 2; Length 325;
Best Local Similarity 28.3%; Pred. No. 1.6e-10;
Matches 89; Conservative 71; Mismatches 124; Indels 30; Gaps 8;
Qy 6 RGLKILSVIGLLFVLIIATAACGNSSNS-----SKESKDGVEIKHEGGTTKPKHPKRV 61
Db 7 RYIKLFILPLTLV--NSCIYINPENSGLITTRIKTSECRLLKHPGLGSCVPVKQVR 64
Qy 62 VLVYSFVDALVALDVKPVGIADNKKRIIKPLDKIGKTVSGTRKQPNLEISKLKP 121
Db 65 IALDETSWEALLADLKPATQAQNIAGSIQKLGKABGIVSLGKQGPNIKMWQLNP 124
Qy 122 DLIADNNRHGIYKDLNKKIAPTIELKSFQGDYNNI--DAFKTISKALGKEEGKKRLE 179
Db 125 DLILGFSVSAEQ-YKLFSQLAPTVTL-----DYIQFQKDALSRITAIIDKSEQAKILLE 178
Qy 180 BHDKKIEYKKEITMDKNQKVLPAVAKSGLLAHPNSNSYVGQFSLQSLGFEKALSDVV--- 236
Db 179 QYQQRVKELRTFINYNLKEKTVSVSRFYAG-----NQVPEFRTKYSFPGLSLTEVRIP 231
Qy 237 ----TKGLSKYLKGPYLQMTETLSQVNPFRMFMTNKAASNEPSLKELEKDPVWKILNA 292
Db 232 VPENQNLTTNENQPLVSVSLERLELDADVLFVALDPGA--ESFQKYQNTPLWQKLDV 289
Qy 293 VKNQRVDILRDLM 306
Db 290 VKNQRVVAVDGYW 303

RESULT 13

AG2129
iron(III) dicitrate-binding periplasmic protein [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AG2129
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AG2129
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-330 <KUR>
A;Cross-references: UNIPROT:O8YTX3; GB:BA000019; PIDN:BA74289.1; PID:gl7131683; GSPDB:C
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr2590

Query Match 17.5%; Score 295.5; DB 2; Length 330;
Best Local Similarity 27.6%; Pred. No. 1e-09;
Matches 93; Conservative 77; Mismatches 130; Indels 37; Gaps 10;
Qy 8 LKILSVIGLLFVLIIATAACGNSSNSKESKDGVE----IKHEGGTTKPKHPKRVV 63
Db 10 LLLLSVCTFLV-----TGCNSNVSKLPNQNTSNTTNCRVINHEAGKTQICGPKKVA 65
Qy 64 LEYSFVDALVALDVKPVGIAD-----DNKNRIIKPLADKI-GKTVSGTRKQPNL 113
Db 66 LSPVPLDMMALGVQAGVAEVDLLNSKVFORPKQ-IPYLDGRITSQPMNVGDRGNPSL 124
Qy 114 BEISKLPDIIADNNRHGIYKDLNKKIAPTIELKSFQGDYNNI--DAFKTISKALGKEE 173
Db 125 ESSLQLKPDILIGEASYNETETQIILNKIAPTILISHRDGITPAWQQTISITAAQLGREDX 184
Qy 174 GKRLREEDKKIEEYKKEIT-MDKNQKVLPAVAKSGL-----LAHPSNSYVGQFSLQSLG 228
Db 185 VPTVIAEYQKLSQAKTALPTAQHQELL--LLAFRGITLTSTFTGAETFFAGLLQDLGF 242
Qy 229 KEALSDDDVTGKLSKYLKGPYLQMTETLSQVNPFRMFMTNKAASNEPSLKELEKDPVWK 288

Db 243 K-LVSPESTE-----YEVALEVLPRKPSGLIVMPSGNNNIENAKRQWQNPILO 291
Qy 289 KLNKVNQRVDILDRDLWARSRGLISSSEMAKELVEL 325
Db 292 SISAHTNRIYFIDYQLASRIRGPIAAELFVNQVRQL 328
RESULT 14
S74458
iron(III) dicitrate-binding periplasmic protein fecB-2a - Synechocystis sp. (strain PCC
N;Alternate names: protein slr1491
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S74458
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-330 <KAN>
A;Cross-references: UNIPROT:P72610; EMBL:D90899; GB:AB001339; NID:gl651650; PIDN:BAAL661
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Genetics:
A;Gene: fecB 2a
C;Superfamily: iron(III) dicitrate transport protein

Query Match 17.3%; Score 292.5; DB 2; Length 330;
Best Local Similarity 28.4%; Pred. No. 1.5e-09;
Matches 85; Conservative 68; Mismatches 119; Indels 27; Gaps 9;

Qy 11 LSVIGLLFVLIIATAACGNSSNS---SSKESKDGVIKHEGGTTKPKHPKRVVVLVYS 67
Db 13 LTIITIVFF-----SACVGSISQNDQSTELLSDVCRIVEHSLGKTCVPLEPRRVVALDGA 68
Qy 68 FVDALVALDVKPVGIADNKKRIIKPLDKIGKTVSGTRKQPNLEISKLKPLDIAD 127
Db 69 TVGNLLALGMMVAGVA-----SNLLPEITRLIPNVPRLGQSSQINLETALVLQDPLIIGA 123
Qy 128 NNRHGIYKDLNKKIAPTIELK-SFGDYNENIDAFKTIKALGKEEGKKLEBHDKKIE 186
Db 124 VWEMKGIYNKLSAIAPTVAFEMQTPADWQR---PFRFDGQVLGLETAQEKVLEQYQMRVN 180
Qy 187 EYKKEITMDKNQKVLPAVAKSGLLA-HPSNSYVGQFSLQSLGFEKALSDDV-TKGLSKYL 244
Db 181 KLRQVSDSPQLQISLVRIRAESGQNSLYLKNCFGAILADLGFARPPSQDQGTDPQPPA 240
Qy 245 KGPYLQMTETLSQVNPFRMFIMT-----NKASSNEPSLKELEKDPVWKILNAVNQRV 298
Db 241 K-----SISRESMTADGVDVIFLFTFGHTPPQIAAAEAQLERLDTDPLWQSLGAVQKNRV 295

RESULT 15

AC2421
iron(III) dicitrate-binding periplasmic protein [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AC2421
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AC2421
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-319 <KUR>

Db 129 ILGFSYSSQD-YKLLSQIAPTAYFNYYTETDWWKTL---LEVAEYVNNKTKRAEKLLEEQH 184

C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

5

C;Accession: A99846
R;Kuroda, M.; Ohta,
ma, A.; Mizutani-Ui,
C.; Shiba T.
T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogu
F;Kuroda, M.; Ohta,
ma, A.; Mizutani-Ui,
C.; Shiba T.
T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogu
F;Kuroda, M.; Ohta,
ma, A.; Mizutani-Ui,
C.; Shiba T.
T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogu

RESULT 21
AE2074
ferrichro

C:Species: Nostoc sp. PCC 7120
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C:Accession: AB2074
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena sp. strain PCC 7120
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AB2074
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-361 <KUR>
 A:Cross-references: UNIPROT:Q8YU34; GB:BA000019; PIDN:BA073846.1; PID:g17131238; GSPDB:G
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: all2147
 C:Superfamily: ferrichrome-iron transport protein fecB

Query Match 16.4%; Score 278; DB 2; Length 361;
 Best Local Similarity 29.3%; Pred. No. 1.1e-08;
 Matches 90; Conservative 63; Mismatches 124; Indels 30; Gaps 13;
 QY 9 KILSVIGLFLVLIATAACGNSSSSSKSSKSGDVE-----IKHEEGTTKVPKPKRV 62
 DB 40 KLPFLPTIVAVVAAC-NNTSQLEQRETSATLSPIETKTISHALGKVKIPLKQPV 98
 QY 63 VLEYSFV-DALVALDVKPGVI--ADNKKQRIIPLDKIGKYSVGT-RKQPNLEEISK 118
 DB 99 VLEENILDSVLGALGKPGVMYQDCENFRGIP-SDLLADVPVGVNIGNQPSLEKILS 157
 QY 119 LKPDLLIADNRRHGKYGKDLNKIAPTIELKSPDG--DYNENIDAKFTISKALGKEEGKK 176
 DB 158 LKPDLLTG-LTWLKSYSKILSSIAPTV-LIDPFSMYDFKERL---RYVAQLGKSDRAE 212
 QY 177 RLEEHDKTIEEYKKTMDKNQKVPAAKSGLLAHPNSYVGQFLSQGFKALSDDV 236
 DB 213 LITQVNRITLRLQGLGKQLEKTSIVI-----YLAGSADIFYRPPDPLAYGQILSDAG 267
 QY 237 TKGSLKYLGPVQMNNTETLSQVNPFRMFMNTKAS-----SNEPSLEKLEKDPVWKILN 291
 DB 268 LRLIQNQKQRELTLISIEVLPKYDADILFIMTEHLTRDFKEANPEVLSPLOK-PIWSNLK 326
 QY 292 AVKNQVR 298
 DB 327 AVKTKQI 333

RESULT 22
 AD2085
 iron(III) dicitrate ABC transporter permease protein fecC [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp. PCC 7120
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C:Accession: AD2085
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena sp. strain PCC 7120
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AD2085
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-319 <KUR>
 A:Cross-references: UNIPROT:Q8YU06; GB:BA000019; PIDN:BA073934.1; PID:g17131326; GSPDB:G
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: fecC
 C:Superfamily: ferrichrome-iron transport protein fecB

Query Match 16.4%; Score 277.5; DB 2; Length 319;
 Best Local Similarity 25.1%; Pred. No. 1e-08;
 Matches 82; Conservative 72; Mismatches 138; Indels 35; Gaps 9;

QY 17 LFLVLIATAACGNSSSSSKSSKSGDVE-----IKHEEGTTKVPKPKRVVLEYSFVDALV 73
 DB 11 LFLVIAISCLISISQSSLPDKISSNLECRLINHELGETCVPLNQRRIIVTQVDAEDVI 70
 QY 74 ALDVPKPGIADNKKQRIIPLDKIGKYSVGT-RKQPNLEEISKLPDILLIADNRRHGK 133
 DB 71 ALGLKPMGAPDPTTVASKSFLSKMSGKINYGKEDQFNLEKIILNLPDILLIISLYGNSA 130
 QY 134 IYKDLNKIAPTIELKSPDGYNENIDAKFTISKALGKEEGKKRLEEHDKKIEEYKKEIT 193
 DB 131 NYQLFSKIAPTQPKYHAKWQ---TSFRQIGEVLGKAEQENLLTEYKQRLTKLRTVLN 187
 QY 194 MDKNQKVPAAKSGLLAHPNSYVGQFLSQGFKALSDDVTKGLS-----KYLKGP 247
 DB 188 NOTNQLKVSYSRPHGV-----QLPEFRSQSFPGSILQEV--GISMPNAQRLIKTP 238
 QY 248 ---YLQNTETLSQVNPFRMFMNTKASSNEPSLKEI-----EKDPVWKILNAVKNQVRDI 300
 DB 239 DDTLIMLNLERIDLDADVLFAV-----DPGARELFQKYONTRLMQTLNVAQNQRVYS 292
 QY 301 LDRDLWARSRLGISSEEMAKELVELSK 327
 DB 293 VDSYSWIFG-SILSANAIVDDLFKYLK 318

RESULT 23

AD2079
 iron(III) dicitrate-binding protein of ABC transporter alt2186 [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp. PCC 7120
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C:Accession: AD2079
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena sp. strain PCC 7120
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AD2079
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-357 <KUR>
 A:Cross-references: UNIPROT:Q8YU25; GB:BA000019; PIDN:BA073885.1; PID:g17131277; GSPDB:G
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: alt2186

Query Match 16.4%; Score 277.5; DB 2; Length 357;
 Best Local Similarity 27.3%; Pred. No. 1.2e-08;
 Matches 98; Conservative 76; Mismatches 134; Indels 51; Gaps 15;

QY 5 VRGLKILSVIGLFLVLIATA-----ACGNSSSSSKSSKSGDVE---IKHEEGTTKVPKH 57
 DB 5 LKSLFVRNRLKLLLAIFSSFMIIACGDRLTQYTVVSQKISPSSTECQVLOHQVGETQVCGQ 64
 QY 58 PKRVVLEYSFVDALVALDVKPGVIAD-----DNKKQRIIPLDKI-KGYTSVGT 107
 DB 65 POKI VALGNVLEILLALDVQPGADHVPLHOGVNDYPSQQ-IPVLGRVGTGELINLGL 123
 QY 108 RKQPNLEEISKLPDILLIADNRRHGKYGKDLNKIAPTIELKSPDGYNENIDAKFTISKA 167
 DB 124 AVTPSIEALLKIKPDLLIIG-NELNRSQVEMLSKIAPTLLFFRFDQPN-----LLAISKI 177
 QY 168 LKHEEGKKRLEEHDKKIEEYKKEIT--MDKNQKVPAAKSGL---LAHP-SNSYVQ 221
 DB 178 LGRSEKAKKILADSDQALIAQKLSLDIVANHPQVLMVLSLQGSISEQLRFRFRSDSCDS 237
 QY 222 FLSQGLQKEALSDDVTGKLSKYLKGPVQMNNTETLSQVNPFRMFMNTKASSN----- 274
 DB 238 LVTNLGFK-LMSPSSLVLPADKSVPPP---LSIETLIPQFNSADSIILGAFNSQVQLSDDK 293
 QY 275 ---EPLSLEKE-----KDPVWKILNAVKNQVRDILDRDLWARSRLISSE-----EMAKELY 323

Db 294 YFEPQMOKLQKQMSNAIAQSLTASKEGRVYFIPAYLCGLPGPIGTGLYLQELQKQL 352

RESULT 24
I39842
iron-uptake system (binding protein) feuD - Bacillus subtilis
C/Species: Bacillus subtilis
C/Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 09-Jul-2004
C/Accession: I39842; E69621
Biochim. Biophys. Acta 1186, 27-34, 1994
A/Title: Isolation of Tn917 insertion mutants of Bacillus subtilis that are resistant
A/Reference number: I39839; MUID:94281248; PMID:8011666
A/Accession: I39842
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-317 <RES>
A/Cross-references: UNIPROT:P40409; GB:L19954; MID:g439454; PIDN:AAA64354.1; PID:g439458
R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A/Authors: Schleich, S.; Schroeter, R.; Scoffone, E.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terptrak, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yanane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A/Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A/Reference number: A69580; MUID:98044033; PMID:9384377
A/Accession: E69621
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-317 <KUN>
A/Cross-references: GB:Z99104; GB:AL009126; MID:g2632267; PIDN:CAB11939.1; PID:g2632430
A/Experimental source: strain 168
C/Genetics:
A/Gene: feuD
C/Superfamily: iron(III) dicitrate transport protein

Query Match 16.4%; Score 276.5; DB 2; Length 317;
Best Local Similarity 26.5%; Pred. NO. 1.2e-08;
Matches 86; Conservative 72; Mismatches 136; Indels 31; Gaps 12;
Qy 9 KILSVIGLLFVLIATAACGNSSSSSK-----ESSKQGVIEKHBEGLTKVPKPKRVVVL 64
Db 3 KISUTLLILLALLTAACGSKNSETASKASGTASEKKIEYLDKTYEVVPTDKIAITGS 62
Qy 65 EYSFVDALVALDVKPVG-IADNKKNIIRIKPLRDKIGKTYSVGTRKQPNLBEISKLPDL 123
Db 63 VESMEDAKL-LDVHPQGAISFGKFPDMFKITDKA---EPGEGWEPNIEKLEMKPDV 118
Qy 124 IADNNRHGIYKOLNKIAPTIELKSPGDYNNENIDAFKTSKALGKEEGKRLLEHDK 183
Db 119 ILASTKFPPEKTLQKISTAGTTPVSHISSNKKENN---MLLAQLTGKSKAKKIADYEQ 175
Qy 184 KIEYKKEIT-WDKNQKVLPAVAASGLLAHPSNYSVQFL-SQLGFEALSDVTKGLS 241
Db 176 DLKEIKTKINDAKDSKALVTRIRGNIIYIPEQVYFNSLTLYGDLGLK-APNEVAAKA 233
Qy 242 KYLGKPYLQMNTEITLSQVNPFRMFI-MTNKASSNEP-SLKELEKDPVWKKLNAVKNQRVD 299
Db 234 QELS-----SLEKLSNNPDHIFVQPSDDENADKPDALKLEKNPIWKSILKAVKEDHY 287
Qy 300 ILDRD-----LWASRGLISSEE 317
Db 288 VNSVDPLAAGGTAWSKVRFLKAAAE 312

Db 294 YFEPQMOKLQKQMSNAIAQSLTASKEGRVYFIPAYLCGLPGPIGTGLYLQELQKQL 352

RESULT 25
S32930
ferrichrome ABC transporter fhud precursor - Bacillus subtilis
N/Alternate names: ferrichrome-binding protein fhud
C/Species: Bacillus subtilis
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C/Accession: S32930; C69622
R/Schneider, R.; Hantke, K.
Mol. Microbiol. 8, 111-121, 1993
A/Title: Iron-hydroxamate uptake systems in Bacillus subtilis: identification of a lipop
A/Reference number: S32930; MUID:93268086; PMID:8388528
A/Accession: S32930
A/Molecule type: DNA
A/Residues: 1-315 <SCH>
A/Cross-references: UNIPROT:P37580; GB:M87283; MID:g289271; PIDN:AAA22424.1; PID:g289272
R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A/Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A/Authors: Schleich, S.; Schroeter, R.; Scoffone, E.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terptrak, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yanane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A/Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A/Reference number: A69580; MUID:98044033; PMID:9384377
A/Accession: C69622
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-315 <KUN>
A/Cross-references: GB:Z99121; GB:AL009126; MID:g2635827; PIDN:CAB15338.1; PID:g2635846
A/Experimental source: strain 168
C/Genetics:
A/Gene: fhud
C/Superfamily: Bacillus subtilis ferrichrome ABC transporter fhud
F/1-23/Domain: signal sequence #status predicted <SIG>
F/24-315/Product: fhud protein #status predicted <MAT>

Query Match 16.3%; Score 275; DB 2; Length 315;
Best Local Similarity 32.1%; Pred. No. 1.4e-08;
Matches 97; Conservative 42; Mismatches 121; Indels 42; Gaps 13;
Qy 17 LFVLIATAACGNSSSSSKSSKDGVE---IKHEEGTTKVPKPKRVVVLVEYSFVDALV 73
Db 15 LLLIALLAACGNSSSKGSASDSK-GAETFTYKAENGNNVKIPKPKRVVWADGYGYPK 73
Qy 74 ALDVKPGIADNKKNIIRIKPLR---DKIGKTYSVGTRKQPNLBEISKLPDLIADNNR 130
Db 74 TLGINVVGAPNVFKNPYKGTNGVENIGDGTG-----EKVIDNLPDLIIVWTQ 125
Qy 131 HKGIYKOLNKIAPTIELKSPGDYNNENIDAFKTSKALGKEEGKRLLEHDKIEEYK 190
Db 126 GADI-KKLEKIAPIVAVKY---DKLDNIEQLEKFAKMTGTEDKAKWLAKDKRVAAK- 180
Qy 191 EITMDKQKQVLPAVAASGLLAHPSNYSVQFLSQLG-----FKEALSDVTKGLSKYL 244
Db 181 -----TKIKKAVGDKTISIMQNGKDIYVFGKDFRGGSGIYKD-LGLQATK-LTKRK 231
Qy 245 ---KGP-YLQNMTEITLSQVNPFRMFI-MTNKASSNEP-SLKELEKDPVWKKLNAVKNQRVDI 300
Db 232 AIDQPGYTSISLEKLPDFAGDYIFAGPWQSGDGGVFE---SSIWKNLNAVKNGHVYK 288
Qy 301 LD 302
Db 289 MD 290

RESULT 26

AG2075
iron(III) dicitrate-binding periplasmic protein of ABC transporter all2157 [imported] -
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AG2075
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AG2075
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-343 <KUR>
A:Cross-references: UNIPROT:Q8YV24; GB:BA000019; PIDN:BA073856.1; PID:gl7131248; GSPDB:G
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all2157
C:Superfamily: ferrichrome-iron transport protein fecB
Query Match 16.2%; Score 274; DB 2; Length 343;
Best Local Similarity 26.5%; Pred. No. 1.8e-08;
Matches 91; Conservative 80; Mismatches 122; Indels 50; Gaps 14;
QY 6 RGLKILS-----VIGLFLVLIATAACGNSS-----SNSKSSKSGVVEIKHEGGTTKV 54
DB 17 KSMKIISRFVFLFLGILIFTAIWVSSKSFYEAATNSIPQNAKCRV-VQHVQGETCI 75
QY 55 PKHPRVVLVLEFVFDVALVDKPVG-----IADD-----NKNRIIPLRDKIGKYS 104
DB 76 PLKQRIIVLDFNSFAALALDTPATITWITEIDDFYFOQKAGVILRSSG----- 131
QY 105 VGTROPNLEESKLPDLIIA-DNNRHGIYKDLNKAIPTELKSF-DGYNENIDAF 161
DB 132 -----QINLEKLVLPDIIIVISHPGFIYKASQIAPTIVLPWIETRGWKKH---- 183
QY 162 KTISKALGEEGKKLEHDKKIEYKEITWMDKNQ-KVLPAVAAKSGLLAHPNSVYG 220
DB 184 QOTARILNRETSIQIYNYNQVRVQKRLNNHRIKHVSFAVAGQVITRQKSPAG 243
QY 221 QFLSQLGFKALSDVDVTKGLSKYLKGPY-LQWNTETLSQVNPMPIMTNKASSNEPSLK 279
DB 244 GILHDIGILNPFAE-----SGDYDLPLESELLPNIDSDILFIAPLR-KDDYSVIK 293
QY 280 ELEKDPVWKLNAVKNQRVDILDRDLWARSGLISSENAKEL 322
DB 294 KLOQKPLWKLKAVQQNQYIVDFSYW-RGLNMLAAYAMLDL 335
RESULT 27
T36412
probable iron-siderophore binding lipoprotein - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T36412
R:Saunders, D.C.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1999
A:Reference number: Z21606
A:Accession: T36412
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-350 <SAU>
A:Cross-references: UNIPROT:Q9RK12; EMBL:AL109974; PIDN:CAB53324.1; GSPDB:GN00070; SCOE
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOE:SCF34.13c
C:Superfamily: iron(III) dicitrate transport protein
Query Match 16.2%; Score 274; DB 2; Length 350;
Best Local Similarity 29.9%; Pred. No. 1.8e-08;
Matches 95; Conservative 56; Mismatches 113; Indels 54; Gaps 14;

QY 14 IGLFLVLIATAACGNSSSNSKSSKDG-----VEIKHEGGTTKVPHKRVVVL 64
DB 25 VGLL-----AGCG--SDSDPADPAGGTTAAAGAFPTVEHAFGTTKIDKAPRVVSV 76
QY 65 EYSFVDALVALDVKPVGAGTADNKNRIIKP-----LRDKIGKYSVGTGRKQ-----PN 112
DB 77 GTTDDQTVLAFGIRKPVGMVDQYNPAGOSPINTQWPVKDKMGDTKPEVVMKNGDTGN 136
QY 113 LEEISKLPDLIIADNNR-HKGIYKDLNKAIPTI-----ELKSPDGDYNNENIDAFKTSK 166
DB 137 FEKIAALRPDLIVAYSEIDQAAAYDKLSKIAPTGTGRTKGEKPPSAPWQDNA---LHIAK 193
QY 167 ALGKEEGKKLEHDKKIEYKKEIT--MDKNQKVLPAVAAKSGLLAHPNSVYQGFLS 224
DB 194 ALGKAEGERKVVADIQGLDAKQAHPFAKDTAVVLSWY--KDSVAPPTSTDVGRGLVT 251
QY 225 QLGPKE-EALSDDVTGKSLKYLKGPYLOMNTETLSQVNPMPIMTNKASSNEPSLKELK 283
DB 252 GIGFKYTEIDKVG-----DFTTTLSPERVDLVDVRFVINDKA--DQDALKKFE- 302
QY 284 DPVWKKLNAVKNQRVDIL 301
DB 303 --LFTNLDAVKNKVSYL 318
RESULT 28
AB2130
iron(III) dicitrate-binding periplasmic protein [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AB2130
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AB2130
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-333 <KUR>
A:Cross-references: UNIPROT:Q8YTX0; GB:BA000019; PIDN:BA074292.1; PID:gl7131686; GSPDB:G
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr2593
C:Superfamily: ferrichrome-iron transport protein fecB
Query Match 16.2%; Score 273; DB 2; Length 333;
Best Local Similarity 28.3%; Pred. No. 1.9e-08;
Matches 90; Conservative 65; Mismatches 127; Indels 36; Gaps 15;
QY 8 LKILSVIGLFLVLIATAACGNSSSNSKSSKDGVE-----IKHEGGTTKVPHKRVV 63
DB 9 LQIIFITLVFVLV--TGCYSHPNQIQSANSKPTVSECRILKHPGETCIPHPORI 66
QY 64 LEYSFVDALVALDVKPVGIA--DDNKNRIIPLRDKIGKYS--VGTGRQNPNEISKL 119
DB 67 MDEDILEITVALDLKPIAVAINFDEWSSR-EKQLWQKABGIDSVIGNHGILNLEKMLLL 125
QY 120 KPDLI--TADNNRHGIYKDLNKAIPTELKSFDDGYNENI--DAFKTISKALGKEEGK 175
DB 126 KPDILGLAESTDRKS-YELFSQIAPTIV-----DYAQTAWRDVLLRGNITIGTEQOQ 179
QY 176 KLEBHDKKIEYKKEITWMDKNQKVLPAVAAKSGLLAHPNSVYQGFLSGLQ--FK 229
DB 180 KLIAEFQRIEKLEKRVIVKNKGLTKLISVVRVNYFHKTIETFRSFPGLSLLVELGLSP 239
QY 230 EALSDDVTG-KLSKYLKGPYLOMNTETLSQVNPMPIMTNKASSNEPSLKELKDPVVK 288
DB 240 EKQNIPTSPG-----PFVFPASLERLDLADVMFVTL--AGGEENFKKFOASPLWQ 291
QY 289 KLNNAVKNQRVDILDRDLW 306
||| ||||| | | |

Query Match	15.7%;	Score 265;	DB 2;	Length 313;
Best Local Similarity	29.2%;	Pred. No. 5e-08;		
Matches	93;	Conservative	53;	Mismatches 116; Indels 56; Gaps 15;
Qy	9	KILSVIGLLFVLITATAACGNSSSSSSKSSKDGVEIKHKEEGTKVPKPKPRVVVLVYSF	68	
Db	3	KVLTTLIAAGALLTLAACSNSSSSKA-TSSDDKVTFLHALNGDVKVPKPKRIAVQNP-	60	
Qy	69	VDALVALDVKPVG-----IADNKNRIIKPLRDKIGKTVSGTKQNRLSEISK	118	
Db	61	-DEVASLGANVVGTDSWAFNTFLSKDQKQNV-----DLGAPKF-NMEKLIA	106	
Qy	119	LKPDLLIADNNRHKGYIKDLNKIAPT--IELKSPDGDYNNENIDAFKTIISKALGKEEKGK	176	
Db	107	QNFDLLIITVD--KQQVADYQKVAPTVLVNKKLSG-MNKSLDYF---AKLLNREDEKAS	159	
Qy	177	RLEHDKKIBEYKKEITMDKNQKVLPAVAAKSGLLAHPNSYV-----GQFLSQ-LG	227	
Db	160	FLNFKFEADQKQEKLA---KVDVVKPAKNTISILELQGDVKYAPGDNFARGGQALTTGLG	216	
Qy	228	FKEALS-DDVTKGLSKYLKGPYLQMNTEITLSQVNPFRFMITMKASNEPSLKELEKDPV	286	
Db	217	FKSSYMWATSKGVG-----YAVNTEISLGEFDADYIFT--DFADKDKQEOYAAKKNPA	268	

QY 287 WKILNAVKNQVDLDRD 304
 | : | | | : | : |
Db 269 WQNLKAVKEGHVITMDYD 286

RESULT 31
AG1319
ferrichrome binding protein homolog lmo1959 [imported] - Listeria monocytogenes
C/Species: Listeria monocytogenes
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C/Accession: AG1319
R/Glaser, P.; Fraigneul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.;
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.;
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournan,
ok, C.; Schluteter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.;
A/Title: Comparative genomics of Listeria species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669

	Query Match	15.7%	Score 264.5	DB 2	Length 313
Best Local Similarity	26.0%	Pred No. 5.4e-08			
Matches 82	Conservative	68	Mismatches 114	Indels 51	Gaps 13
Qy	10	ILSVIGLLFVLIIATAACGNSSSSSKSSK-----DGEVIKHBEGTTKVPKPKRVVVL	64		
Db	5	IILVSMLFIAVLVTACDGNKSAGSEQVEMRTYTMWANGKV-----EIAHPKRIIVAS	57		
Qy	65	EYSFVDALVALDVKPVGIADNKNRIIKLPDKIGKVTYVGRKQP--NLREISKLPDIL	123		
Db	58	EY--LGRNIVLIGMKPVGARQKMNPFLL-----GKVDGADTGDPPVSAKVAELKKDIL	109		
Qy	124	IIADNNRHKGTYKDLNKIAPTIEL-----KSFDDGYNENIDAFKTTISKALGCKEGKKRL	178		
Db	110	IIVSNDDE---FEAMSKIAPTVLPYATSKNVBEEDVRQ-----IADLVGEKKAGEAWL	159		
Qy	179	BEHKKTIEYKKEIT--MDKNQKV-LPVAASKGILLAHPSNYSYVGQFL---SOLGFKCAL	232		
Db	160	DKFHQKAKESRAKLAGLDPNETGTIVVEVDKDPYVMQGNMGRGGOALYNALQLKAPAKI	219		

```
enterochelin uptake periplasmic binding protein Cj1355 [imported] - Campylobacter jejuni
C:Species: Campylobacter jejuni
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: H81279
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chiller,
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrell
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hy
A:Reference number: A81250; MUID:20150912; PMID:10688204
A:Accession: H81279
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-330 <PAR>
A:Cross-references: UNIPROT:Q9PMU4; GB:AL139078; GB:AL111168; NID:g6968723; PIDN:CAB7376
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: ceuE; Cj1355

Query Match 15.4%; Score 261; DB 2; Length 330;
Best Local Similarity 26.3%; Pred. NO. 9e-08;
Matches 88; Conservative 72; Mismatches 123; Indels 52; Gaps 14;

QY 18 FVLATACGNNSSNSSKSSK-----DGVETKHEEGTTKVPKPKR 60
DB 13 FLSLITACNSNENNASSSTTKNTATVKVLPISMDEGDSFLVKDSLGENKIPKNPSK 72
QY 61 VVLEYSFVDALVALDV--KPVGIADNNKRIIKPLDKIKGKTVSGTRKPNLEEISK 118
DB 73 VVLDLGLDTFDALKUNDKVVGVPAKNLP-KYLQOFKVK-----PSVGGVQVDFEALNA 127
QY 119 LKPOLIITADNNRHGIYKDLNKIAPTIELKSPDGDYNN--IDAFK---TISKALGKEE 173
DB 128 LKPOLIITISGRQSK-FYDKLKEIAPT-----FVGLDANFLSSPENNVLVAKLYGLEKE 182
QY 174 GKXLEHDKKIEYKEITMDKQKVPVAAKSGLLAHSNSVVGFLSOLGFEALS 233
DB 183 ALEKISDIKNEIEKAIS--IVDEKKAIIILTSNKISAFQPSRFGIHDVLGI-NAVD 239
QY 234 DDV---TKGLSKYLPVQLQNTETLSOVNPERFMIM-TNKASSNEPSLKEKDPVWKK 289
DB 240 ENIKVGTGHS-----INSEFILEKNPDYIFVDRNVILGNKERAQGLDNLVAK 290
QY 290 LNAVKNQKRVLDLDRDW--ARSRLGISSEEMAKEL 322
DB 291 TKAAQNKKIILYDPEYWLASNGLESKLTWILEI 325

RESULT 34
S74441
iron(III) dicitrate transport system permease protein fecB - Synechocystis sp. (strain
N:Alternate names: protein slr1319
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S74441
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasu:
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocysti
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S74441
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-315 <KAN>
A:Cross-references: UNIPROT:P72593; EMBL:D90899; GB:AB001339; NID:g1651650; PIDN:BAAL65.
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Gene: fecB
C:Superfamily: iron(III) dicitrate transport protein
C:Keywords: iron transport

Query Match 15.4%; Score 260; DB 1; Length 315;

enterochelin uptake periplasmic binding protein Cj1355 [imported] - Campylobacter jejuni
C:Species: Campylobacter jejuni
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: H81279
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chiller,
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrell
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hy
A:Reference number: A81250; MUID:20150912; PMID:10688204
A:Accession: H81279
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-330 <PAR>
A:Cross-references: UNIPROT:Q9PMU4; GB:AL139078; GB:AL111168; NID:g6968723; PIDN:CAB7376
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: ceuE; Cj1355

Query Match 15.4%; Score 261; DB 2; Length 330;
Best Local Similarity 26.3%; Pred. NO. 9e-08;
Matches 88; Conservative 72; Mismatches 123; Indels 52; Gaps 14;

QY 18 FVLATACGNNSSNSSKSSK-----DGVETKHEEGTTKVPKPKR 60
DB 13 FLSLITACNSNENNASSSTTKNTATVKVLPISMDEGDSFLVKDSLGENKIPKNPSK 72
QY 61 VVLEYSFVDALVALDV--KPVGIADNNKRIIKPLDKIKGKTVSGTRKPNLEEISK 118
DB 73 VVLDLGLDTFDALKUNDKVVGVPAKNLP-KYLQOFKVK-----PSVGGVQVDFEALNA 127
QY 119 LKPOLIITADNNRHGIYKDLNKIAPTIELKSPDGDYNN--IDAFK---TISKALGKEE 173
DB 128 LKPOLIITISGRQSK-FYDKLKEIAPT-----FVGLDANFLSSPENNVLVAKLYGLEKE 182
QY 174 GKXLEHDKKIEYKEITMDKQKVPVAAKSGLLAHSNSVVGFLSOLGFEALS 233
DB 183 ALEKISDIKNEIEKAIS--IVDEKKAIIILTSNKISAFQPSRFGIHDVLGI-NAVD 239
QY 234 DDV---TKGLSKYLPVQLQNTETLSOVNPERFMIM-TNKASSNEPSLKEKDPVWKK 289
DB 240 ENIKVGTGHS-----INSEFILEKNPDYIFVDRNVILGNKERAQGLDNLVAK 290
QY 290 LNAVKNQKRVLDLDRDW--ARSRLGISSEEMAKEL 322
DB 291 TKAAQNKKIILYDPEYWLASNGLESKLTWILEI 325

RESULT 34
S74441
iron(III) dicitrate transport system permease protein fecB - Synechocystis sp. (strain
N:Alternate names: protein slr1319
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S74441
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasu:
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocysti
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S74441
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-315 <KAN>
A:Cross-references: UNIPROT:P72593; EMBL:D90899; GB:AB001339; NID:g1651650; PIDN:BAAL65.
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Gene: fecB
C:Superfamily: iron(III) dicitrate transport protein
C:Keywords: iron transport

Query Match 15.4%; Score 260; DB 1; Length 315;
```

```
enterochelin uptake periplasmic binding protein Cj1355 [imported] - Campylobacter jejuni
C:Species: Campylobacter jejuni
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: H81279
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chiller,
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrell
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hy
A:Reference number: A81250; MUID:20150912; PMID:10688204
A:Accession: H81279
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-330 <PAR>
A:Cross-references: UNIPROT:Q9PMU4; GB:AL139078; GB:AL111168; NID:g6968723; PIDN:CAB7376
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: ceuE; Cj1355

Query Match 15.4%; Score 261; DB 2; Length 330;
Best Local Similarity 26.3%; Pred. NO. 9e-08;
Matches 88; Conservative 72; Mismatches 123; Indels 52; Gaps 14;

QY 18 FVLATACGNNSSNSSKSSK-----DGVETKHEEGTTKVPKPKR 60
DB 13 FLSLITACNSNENNASSSTTKNTATVKVLPISMDEGDSFLVKDSLGENKIPKNPSK 72
QY 61 VVLEYSFVDALVALDV--KPVGIADNNKRIIKPLDKIKGKTVSGTRKPNLEEISK 118
DB 73 VVLDLGLDTFDALKUNDKVVGVPAKNLP-KYLQOFKVK-----PSVGGVQVDFEALNA 127
QY 119 LKPOLIITADNNRHGIYKDLNKIAPTIELKSPDGDYNN--IDAFK---TISKALGKEE 173
DB 128 LKPOLIITISGRQSK-FYDKLKEIAPT-----FVGLDANFLSSPENNVLVAKLYGLEKE 182
QY 174 GKXLEHDKKIEYKEITMDKQKVPVAAKSGLLAHSNSVVGFLSOLGFEALS 233
DB 183 ALEKISDIKNEIEKAIS--IVDEKKAIIILTSNKISAFQPSRFGIHDVLGI-NAVD 239
QY 234 DDV---TKGLSKYLPVQLQNTETLSOVNPERFMIM-TNKASSNEPSLKEKDPVWKK 289
DB 240 ENIKVGTGHS-----INSEFILEKNPDYIFVDRNVILGNKERAQGLDNLVAK 290
QY 290 LNAVKNQKRVLDLDRDW--ARSRLGISSEEMAKEL 322
DB 291 TKAAQNKKIILYDPEYWLASNGLESKLTWILEI 325

RESULT 34
S74441
iron(III) dicitrate transport system permease protein fecB - Synechocystis sp. (strain
N:Alternate names: protein slr1319
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S74441
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasu:
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocysti
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S74441
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-315 <KAN>
A:Cross-references: UNIPROT:P72593; EMBL:D90899; GB:AB001339; NID:g1651650; PIDN:BAAL65.
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Gene: fecB
C:Superfamily: iron(III) dicitrate transport protein
C:Keywords: iron transport

Query Match 15.4%; Score 260; DB 1; Length 315;
```


A: Experimental source: strain PCC 7120
C: Genetics:
A: Gene: alr2597

Query Match 15.0%; Score 254; DB 2; Length 349;
Best Local Similarity 26.5%; Pred. No. 2.4e-07;
Matches 96; Conservative 73; Mismatches 129; Indels 64; Gaps 17;

QY 9 KILSVIGLLFVLIAAACGNNSSSSKES-----SKOGVEIKHEGGTTKPKPKRVVVL 64
DB 13 KYIVLVLTSLI--TACGGDIIDNPLESNVPOSTDCRVHEHVGKTKICGPKKIAVL 70
QY 65 EYFVDALVALDVKPVGIAD-----DNKNRIIKPLRDKIGKYT-----SVGTRKQ 110
DB 71 EPKLSLMLADVQAAADALYNAPQDFNPSQI-----PYLGKYVTSQPINIGDRSS 125
QY 111 PNLEISKLPDLIIADNRHGIYKD---LNKIAPTIELSKFDGYNENIDAFKTIKA 167
DB 126 PSLEALTLLKPLDILGLN-----YQDNQLFSAIAPTTFI---DNEQNWQ-ESIKIVAKA 175
QY 168 LKKEEGKKRLEHDKKIEYKKEI--TMDKNQKVLPAVAKSGLLAHPNSYVG---QF 222
DB 176 LDSQKNIPSITISQOQAKRAQLESIVSTHPRVLNIVC--SQMDYIEVTYNGNTIEI 233
QY 223 LSQLGFEALSDDVTK--GLSKYLKGPYLQMNTEITLSQVNPFRMFT-----NKASSNE 275
DB 234 LEKIGFQPVLPDIERKGLRP-----QITLETLAQDADIVIVNTWVDNNGKSYT 286
QY 276 PSLKELE---KDPVKKLNAVKNQVRDILDRDLWARSGL--ISSEEMAKELVELSKSK 330
DB 287 VFLKELKQKWAKNPILLNSQAWKEGVYVDYTLGNVIGAPIANALILEQLPSLLSNT 346
QY 331 KK 332
DB 347 LK 348

RESULT 38
G95218
hypothetical protein SP1872 [imported] - Streptococcus pneumoniae (strain TIGR4)
C: Species: Streptococcus pneumoniae
C: Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C: Accession: G95218
R: Rettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfel,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A: Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A: Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.
A: Reference number: A95000; MUID: 21357209; PMID: 11463916
A: Accession: G95218
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-321 <KUR>
A: Cross-references: UNIPROT: Q97NX9; GB: AE005672; PIDN: AAK75944.1; PID: g14973375; GSPDB: G
A: Experimental source: strain TIGR4
C: Genetics:
A: Gene: SP1872

Query Match 15.0%; Score 253; DB 2; Length 321;
Best Local Similarity 26.8%; Pred. No. 2.4e-07;
Matches 91; Conservative 66; Mismatches 117; Indels 66; Gaps 14;

QY 20 LIATAACGNNSSSSKESKSGVE--IKHEGTTKVPKPKRVVVLVSFVDALVALDV 77
DB 17 LLLLGACSTNSSTQETSSAPTEVITIKSSDEVKLSKVPEKIVTFDILGAADTIRALGF 76
QY 78 KPVGIADNNKNRIKP-----LRDKIGKTVSVTRKQPNLEISKLKPDIIADNR 130
DB 77 -----EKNIVGMPTKTVPYTKDLVGTGVKVGSKPEPDLAIAALEPDLIIASPT 127
QY 131 HKGIYKDLNKIAPTIELSKFDGYNENIDA-FKTIKALGK--EEGKKRLEHDKKIEE 187
QY 131 HKGIYKDLNKIAPTIELSKFDGYNENIDA-FKTIKALGK--EEGKKRLEHDKKIEE 187

DB 128 QKFVDK-FKEIAPTFLVFOASKDDYWTSTKANIESLASAFGETGTQAKKEELTKLKDKSIOE 186
QY 188 Y-----KKEITMDKNQKVLPAVAKSGLLAHPNSYVQFLSQ-LGFK--BALSDDV 236
DB 187 VATKNESDDKALAILNEGKMAAFGAKSRF-----SFLYQTLKFKPTDTKFEDS 236
QY 237 TKGLSKYLKGPYLQMNTEITLSQVNPFRMFTWKA-----SSNEPSLKE--LEKDPVWK 288
DB 237 RHG-----QEVSFESVKEINPDILFVINRTLAIGDSSNDGVLNIALIAETP--- 284
QY 289 KLNNAVKNQVRDILDRDLWARSGLISSEEMAKELVELSKK 328
DB 285 ---AAKNGKIQLTPDLWYLSGGLESTKLMIEDIQALK 321

RESULT 39

E98082
hypothetical protein fatB [imported] - Streptococcus pneumoniae (strain R6)
C: Species: Streptococcus pneumoniae
C: Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C: Accession: E98082
R: Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.;
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.;
J. P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A: Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaekunas, S.R.;
A: Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A: Reference number: A97872; MUID: 21429245; PMID: 11544234
A: Accession: E98082
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-321 <KUR>
A: Cross-references: UNIPROT: Q8DNJ2; GB: AE007317; PIDN: AAL00490.1; PID: g15459362; GSPDB: G
C: Genetics:
A: Gene: fatB

Query Match 15.0%; Score 253; DB 2; Length 321;
Best Local Similarity 26.8%; Pred. No. 2.4e-07;
Matches 91; Conservative 66; Mismatches 117; Indels 66; Gaps 14;

QY 20 LIATAACGNNSSSSKESKSGVE--IKHEGTTKVPKPKRVVVLVSFVDALVALDV 77
DB 17 LLLLGACSTNSSTQETSSAPTEVITIKSSDEVKLSKVPEKIVTFDILGAADTIRALGF 76
QY 78 KPVGIADNNKNRIKP-----LRDKIGKTVSVTRKQPNLEISKLKPDIIADNR 130
DB 77 -----EKNIVGMPTKTVPYTKDLVGTGVKVGSKPEPDLAIAALEPDLIIASPT 127
QY 131 HKGIYKDLNKIAPTIELSKFDGYNENIDA-FKTIKALGK--EEGKKRLEHDKKIEE 187
DB 128 QKFVDK-FKEIAPTFLVFOASKDDYWTSTKANIESLASAFGETGTQAKKEELAKLDSIOE 186
QY 188 Y-----KKEITMDKNQKVLPAVAKSGLLAHPNSYVQFLSQ-LGFK--EALSDDV 236
DB 187 VATKNESDDKALAILNEGKMAAFGAKSRF-----SFLYQTLKFKPTDTKFEDS 236
QY 237 TKGLSKYLKGPYLQMNTEITLSQVNPFRMFTWKA-----SSNEPSLKE--LEKDPVWK 288
DB 237 RHG-----QEVSFESVKEINPDILFVINRTLAIGDSSNDGVLNIALIAETP--- 284
QY 289 KLNNAVKNQVRDILDRDLWARSGLISSEEMAKELVELSKK 328
DB 285 ---AAKNGKIQLTPDLWYLSGGLESTKLMIEDIQALK 321

RESULT 40

DB3778
hypothetical protein BH1028 [imported] - Bacillus halodurans (strain C-125)
C: Species: Bacillus halodurans
C: Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C: Accession: DB3778
R: Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hir
Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: D83778
A;Status: Preliminary
A;Molecule type: DNA
A;Residues: 1-329 <STO>
A;Cross-references: UNIPROT:O9KE30; GB:AP001510; GB:BA000004; NID:gi0173440; PIDN:BA047
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH1028

Query Match		14.8%	Score 250.5;	DB 2;	Length 329;
Best Local Similarity		25.6%	Pred. No. 3.5e-07;		
Matches		86;	Conservative 71;	Mismatches 134;	Indels 45; Gaps 13;
Qy	10	ILSVIGLLFVLIATAACNNSSN-----SSKESKDGVE-----IKHEEGTT	52		
Db	11	VVGIVSLMW-----ACGADTAPADAEQEAAQDTVEBEMVETEAGGEVTVQHDLGEA	65		
Qy	53	KVPKHPKRVVVLVYSFVDALVALDVPVGIADDNKKRIIKPLRDKIKGYTSVGTAKQPN	112		
Db	66	VIPKHPETVVVDFGTLDLRLGVEVAGVPAQNPAYLSQYBEE---TYANVGTLPFPD	122		
Qy	113	LEEISKLPDLIIADNNRHKGIYKDLNKIAPTIELKSGFDGYNEN-IDAPK---TISKA	167		
Db	123	PETIYDQPDIIII-SGRTSAYDELSDIAPTL----FVGLDTENYLESFRSNMETLGEI	177		
Qy	168	LGKEEGKKRLLEHDKIEEYKEITMDKNQKVLPAVAAKSGLLAHPSNSYVGQFLSQLG	227		
Db	178	FGKETEVEEALASLEASTEIV-QQLAAEKETGLIVLANDGNVSAYGPGSRFGVIHDEFG	236		
Qy	228	FKEALSDDVTKGLSKYLKGPYLQNTETLSQVNPERPMTNKAASSNEPSLKELEKDPVW	287		
Db	237	V-----TPVDENIEVSNHGQ--SISFEYIVKPNPDHLPVIDRGAVVQEGEETNIENELV-	288		
Qy	288	KKLNAVKNQRVDIILDRDLWARS-RGLISSEEMAKEL	322		
Db	289	QQTAYEGNIHYLTPDIWYISGGGLSVSQMIDEM	324		

Search completed: August 25, 2005, 09:27:10
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 09:27:00 ; Search time 163 Seconds
(without alignments)
807.548 Million cell updates/sec

Title: US-10-724-972A-6352
Perfect score: 1690
Sequence: 1 GVESVRGLKILSVIGLLFVL.....BEMAKVELSKDKSKDNK 335

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1767149 seqs, 392926209 residues

Total number of hits satisfying chosen parameters: 1767149

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1690	100.0	335	US-10-724-972A-6352	Sequence 6352, Ap
2	1668	98.7	331	US-10-282-122A-71138	Sequence 71138, A
3	1242	73.5	337	US-09-815-242-12254	Sequence 12254, A
4	1232	72.9	327	US-10-282-122A-43828	Sequence 43828, A
5	1181	69.9	309	US-09-815-242-5459	Sequence 5459, Ap
6	726	43.0	319	US-09-815-242-5249	Sequence 5249, Ap
7	726	43.0	319	US-09-815-242-12122	Sequence 12122, A
8	724.5	42.9	319	US-10-282-122A-41133	Sequence 41133, A
9	724.5	42.9	319	US-10-470-048B-329	Sequence 329, App
10	644	38.1	175	US-10-282-122A-71291	Sequence 71291, A
11	516.5	30.6	298	US-10-282-122A-66890	Sequence 66890, A

12	460	27.2	300	15	US-10-282-122A-56054	Sequence 56054, A
13	460	27.2	302	9	US-09-815-242-10469	Sequence 10469, A
14	460	27.2	302	15	US-10-282-122A-56805	Sequence 56805, A
15	454	26.9	297	15	US-10-282-122A-69420	Sequence 69420, A
16	405	24.0	330	8	US-08-781-986A-5193	Sequence 5193, Ap
17	405	24.0	330	14	US-10-278-946-16	Sequence 16, Appl
18	405	24.0	330	15	US-10-329-624-5193	Sequence 5193, Ap
19	405	24.0	330	17	US-10-470-048B-346	Sequence 346, App
20	405	24.0	330	17	US-10-967-189-16	Sequence 16, Appl
21	391	23.1	324	15	US-10-282-122A-45599	Sequence 45599, A
22	331	19.6	315	9	US-09-738-626-3875	Sequence 3875, Ap
23	327	19.3	321	15	US-10-282-122A-53267	Sequence 53267, A
24	313	18.5	314	15	US-10-282-122A-46283	Sequence 46283, A
25	301	17.8	314	15	US-10-282-122A-45385	Sequence 45385, A
26	300	17.8	314	15	US-10-282-122A-46193	Sequence 46193, A
27	296.5	17.5	351	18	US-10-724-972A-7181	Sequence 7181, Ap
28	296	17.5	342	15	US-10-282-122A-68785	Sequence 68785, A
29	283	16.7	317	9	US-09-071-035-38	Sequence 38, Appl
30	283	16.7	317	14	US-10-206-576-38	Sequence 38, Appl
31	283	16.7	317	17	US-10-912-362-38	Sequence 38, Appl
32	279	16.5	317	9	US-09-815-242-10905	Sequence 10905, A
33	274.5	16.2	315	15	US-10-282-122A-46179	Sequence 46179, A
34	268	15.9	296	9	US-09-071-035-40	Sequence 40, Appl
35	268	15.9	296	14	US-10-206-576-40	Sequence 40, Appl
36	268	15.9	296	17	US-10-912-362-40	Sequence 40, Appl
37	266	15.7	306	9	US-09-738-626-4403	Sequence 4403, App
38	266	15.7	306	15	US-10-282-122A-47311	Sequence 47311, A
39	257.5	15.2	342	15	US-10-474-776-663	Sequence 663, App
40	254	15.0	341	16	US-10-415-478A-12	Sequence 12, Appl
41	254	15.0	341	16	US-10-472-928-2046	Sequence 2046, Ap
42	254	15.0	341	17	US-09-815-242-13454	Sequence 13454, A
43	253	15.0	321	9	US-10-415-478A-9	Sequence 9, Appl
44	253	15.0	321	16	US-10-472-928-3884	Sequence 3884, Ap
45	253	15.0	321	17	US-10-472-928-3884	Sequence 3884, Ap

ALIGNMENTS

RESULT 1
US-10-724-972A-6352
; Sequence 6352, Application US/10724972A
; Publication No. US20040147734A1
; GENERAL INFORMATION:
; APPLICANT: Doucette-Stamm, Lynn
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PATH03-16
; CURRENT APPLICATION NUMBER: US/10/724,972A
; PRIOR FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: 09/450,969
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: 09/134,001
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 7544
; SEQ ID NO 6352
; LENGTH: 335
; TYPE: PRT
; ORGANISM: S.epidermidis
US-10-724-972A-6352

Query Match 100.0%; Score 1690; DB 18; Length 335;
Best Local Similarity 100.0%; Pred. No. 6.8e-107;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVESVRGLKILSVIGLLFVLITACGNNSSNSSKSSKDGVEIKHEGTTKVPKPKR 60
DB 1 GVESVRGLKILSVIGLLFVLITACGNNSSNSSKSSKDGVEIKHEGTTKVPKPKR 60


```
Db 10 TMRGLKTFESILGLI VALFLVAACGNTDSSSKESSTKDTISVKDENGTVKVPKDAKRIW 69
Qy 64 LEYSFVDALVALDVKPGIADNKNRII KPLRDKIGKTVTSVGTQKNLEISKLPDL 123
Db 70 LEYSFADALVALDVKPGIADNKNRII KPVREKIGNYTSVGTQKNLEISKLPDL 129
Qy 124 IADNNRHGIYKOLNKIAPTIELSKFSGDYNENIDAFKTI SKALGKEEGKRLAEHDK 183
Db 130 IADSSRHKGINKELNKIAPTIELSKFSGDYGKQNSFKTI AKALNKEKEGKRLAEHDK 189
Qy 184 KIEEYKEITMDKNOKVLPVAAKSGLLAHPNSVVGQFLSOLGPKFKEALSDDVTKGLSKY 243
Db 190 LIKKYKDEIKFDRNOKVLPVAVAKAGLLAHPNSVYVGQFLNELGPKFNAISDDVTKGLSKY 249
Qy 244 LKGPVLQNTETLSQVNERPMFIMTNKASSNEPSLKELEKDPVWKKNVKNORVDIILDR 303
Db 250 LKGPYLQDLTEHLADLNPERMIITMDNAKOSAEFKLQEDPTWKKLNNAVKNRVDIIVDR 309
Qy 304 DWARSRLISSEEMAKELVELSKKDSK 331
Db 310 DWARSRLISSEEMAKELVELSKKEQK 337
```

```
RESULT 4
US-10-282-122A-43828
; Sequence 43828, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43828
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-282-122A-43828
```

```
Query Match 72.9%; Score 1232; DB 15; Length 327;
Best Local Similarity 71.9%; Pred. No. 9.1e-76;
Matches 235; Conservative 42; Mismatches 50; Indels 0; Gaps 0;
Qy 5 VRGKILSVITGLLFLVIATAACGNSSNSKSSKSGVVEIKHEEGTTKVPKPKRVVVL 64
Db 1 MRGLKTFESILGLI VALFLVAACGNTDSSSKESSTKDTISVKDENGTVKVPKDAKRIW 60
Qy 65 BYSFVDALVALDVKPGIADNKNRII KPLRDKIGKTVTSVGTQKNLEISKLPDL 124
Db 61 EYSFADALVALDVKPGIADNKNRII KPVREKIGNYTSVGTQKNLEISKLPDL 120
Qy 125 IADNNRHGIYKOLNKIAPTIELSKFSGDYNENIDAFKTI SKALGKEEGKRLAEHDK 184
Db 121 IADSSRHKGINKELNKIAPTIELSKFSGDYGKQNSFKTI AKALNKEKEGKRLAEHDK 180
Qy 185 IEEYKEITMDKNOKVLPVAAKSGLLAHPNSVVGQFLSOLGPKFKEALSDDVTKGLSKY 244
Db 181 INKYKDEIKFDRNOKVLPVAVAKAGLLAHPNSVYVGQFLNELGPKFNAISDDVTKGLSKY 240
Qy 245 KGPVLQNTETLSQVNERPMFIMTNKASSNEPSLKELEKDPVWKKNVKNORVDIILDR 304
Db 241 KGPYLQDLTEHLADLNPERMIITMDNAKOSAEFKLQEDPTWKKLNNAVKNRVDIIVDR 300
Qy 305 LWARSRLISSEEMAKELVELSKKDSK 331
Db 301 VWARSRLISSEEMAKELVELSKKEQK 327
```

```
RESULT 5
US-09-815-242-5459
; Sequence 5459, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5459
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5459
Query Match 69.9%; Score 1181; DB 9; Length 309;
Best Local Similarity 72.8%; Pred. No. 2.5e-72;
Matches 225; Conservative 38; Mismatches 46; Indels 0; Gaps 0;
```


QY	126	ADNNRHGIYKDLNKIAPTIELKSPDGYNENIDAFKTIISKALGKEEGKRLBEHDKKI	185
Db	117	ADVSRHKIKSELSKIAPTILVSGTGDYNANIDAFKTVAKAVGKEGKRLBEHDKKI	176
QY	186	BEYKKEITMDNKNQKVLPAVAKSGLLAHPNSYVGQFLSQGLFKALSDDTVTKGLSKYLK	245
Db	177	AEIRKKIEQSTLKTAFAFGISRAGFINNEDTFMGQFLIKMGIOPEVTKDTHVGERKG	236
QY	246	GPYLOMNTETLSQVNPFRMFWTN-KASSNEPSLKELEKDP-VWKKLNQVNRVDILDR	303
Db	237	GPYIYNNEELANINPKVMILATDGTDKNRTKF----IDPAVWKS LKAVKDNKYVDVDR	292
QY	304	DLWARSGLISSEEMAKELVELSKK 328	
Db	293	NKWLKSRGIIASESMAEDLEKIAEK 317	
RESULT 8			
US-10-282-122A-44133			
; Sequence 44133, Application US/10282122A			
; Publication No. US20040029129A1			
; GENERAL INFORMATION:			
; APPLICANT: Wang, Liangsu			
; APPLICANT: Zamudio, Carlos			
; APPLICANT: Malone, Cheryl			
; APPLICANT: Haselbeck, Robert			
; APPLICANT: Ohlsen, Kari			
; APPLICANT: Zyskind, Judith			
; APPLICANT: Wall, Daniel			
; APPLICANT: Trawick, John			
; APPLICANT: Cart, Grant			
; APPLICANT: Yamamoto, Robert			
; APPLICANT: Forsyth, R.			
; APPLICANT: Xu, H.			
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms			
; FILE REFERENCE: ELITRA.034A			
; CURRENT APPLICATION NUMBER: US/10/282,122A			
; CURRENT FILING DATE: 2003-02-20			
; PRIOR APPLICATION NUMBER: 60/191,078			
; PRIOR FILING DATE: 2000-03-21			
; PRIOR APPLICATION NUMBER: 60/206,848			
; PRIOR FILING DATE: 2000-05-23			
; PRIOR APPLICATION NUMBER: 60/207,727			
; PRIOR FILING DATE: 2000-05-26			
; PRIOR APPLICATION NUMBER: 60/230,335			
; PRIOR FILING DATE: 2000-09-06			
; PRIOR APPLICATION NUMBER: 60/230,347			
; PRIOR FILING DATE: 2000-09-09			
; PRIOR APPLICATION NUMBER: 60/242,578			
; PRIOR FILING DATE: 2000-10-23			
; PRIOR APPLICATION NUMBER: 60/253,625			
; PRIOR FILING DATE: 2000-11-27			
; PRIOR APPLICATION NUMBER: 60/257,931			
; PRIOR FILING DATE: 2000-12-22			
; PRIOR APPLICATION NUMBER: 60/267,636			
; PRIOR FILING DATE: 2001-02-09			
; PRIOR APPLICATION NUMBER: 60/269,308			
; PRIOR FILING DATE: 2001-02-16			
; Remaining Prior Application data removed - See File Wrapper or PALM.			
; NUMBER OF SEQ ID NOS: 78614			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 44133			
; LENGTH: 319			
; TYPE: PRT			
; ORGANISM: Staphylococcus aureus			
US-10-282-122A-44133			
Query Match 42.9%; Score 724.5; DB 15; Length 319;			
Best Local Similarity 46.9%; Pred. No. 2.8e-41;			
Matches 149; Conservative 63; Mismatches 95; Indels 11; Gaps 4;			
QY	13	VIGLLFVLATACGNNSSNSSKSSKDGVEIKHEEGTTKVPKPKRVVVLSEYSPDAL	72

Db	9	VVFMILVAVAGCGQKDT-----BEKTEMTTIKDELGTETKIKKPKRVVVLSEYSPDAL	63
QY	73	VALDVKPVGIIADNKNRIIPKLRDKIGKYSVGRKQPNLEIEISKLPDLIIADNNRHK	132
Db	64	AALDNKPVGIADDDGSGTKNITKSVDRKIGAYESVGRPQPNMEVSKLPDLIIADVSRHK	123
QY	133	GIYKDLNKIAPTIELKSPDGYNENIDAFKTIISKALGKEEGKRLBEHDKKI	192
Db	124	KIKSELSKIAPTILVSGTGDYNANIEAFKTVAKAVGKEGKRLBEHDKKI	183
QY	193	TMDKNQKVLPAVAKSGLLAHPNSYVGQFLSQGLFKALSDDTVTKGLSKYLKGYLOWN	252
Db	184	EQSTLKSFAFGISRAGFINNEDTFMGQFLIKMGIOPEVTKDTHVGERKGGPYIYN	243
QY	253	TETLSQVNPFRMFWTN-KASSNEPSLKELEKDP-VWKKLNQVNRVDILDRDLWARSR	310
Db	244	NEELANINPKVMILATDGTDKNRTKF----IDPAVWKS LKAVKDNKYVDVDRNKWLKSR	299
QY	311	GLISSEEMAKELVELSKK 328	
Db	300	GIIASESMAEDLEKIAEK 317	
RESULT 9			
US-10-470-048B-329			
; Sequence 329, Application US/10470048B			
; Publication No. US20050037444A1			
; GENERAL INFORMATION:			
; APPLICANT: MEINKE ET AL.			
; TITLE OF INVENTION: A METHOD FOR IDENTIFICATION, ISOLATION AND PRODUCTION OF			
; FILE REFERENCE: SONN:035US			
; CURRENT APPLICATION NUMBER: US/10/470,048B			
; CURRENT FILING DATE: 2003-07-25			
; NUMBER OF SEQ ID NOS: 603			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 329			
; LENGTH: 319			
; TYPE: PRT			
; ORGANISM: Staphylococcus aureus			
US-10-470-048B-329			
Query Match 42.9%; Score 724.5; DB 17; Length 319;			
Best Local Similarity 46.9%; Pred. No. 2.8e-41;			
Matches 149; Conservative 63; Mismatches 95; Indels 11; Gaps 4;			
QY	13	VIGLLFVLATACGNNSSNSSKSSKDGVEIKHEEGTTKVPKPKRVVVLSEYSPDAL	72
Db	9	VVFMILVAVAGCGQKDT-----BEKTEMTTIKDELGTETKIKKPKRVVVLSEYSPDAL	63
QY	73	VALDVKPVGIIADNKNRIIPKLRDKIGKYSVGRKQPNLEIEISKLPDLIIADNNRHK	132
Db	64	AALDNKPVGIADDDGSGTKNITKSVDRKIGAYESVGRPQPNMEVSKLPDLIIADVSRHK	123
QY	133	GIYKDLNKIAPTIELKSPDGYNENIDAFKTIISKALGKEEGKRLBEHDKKI	192
Db	124	KIKSELSKIAPTILVSGTGDYNANIEAFKTVAKAVGKEGKRLBEHDKKI	183
QY	193	TMDKNQKVLPAVAKSGLLAHPNSYVGQFLSQGLFKALSDDTVTKGLSKYLKGYLOWN	252
Db	184	EQSTLKSFAFGISRAGFINNEDTFMGQFLIKMGIOPEVTKDTHVGERKGGPYIYN	243
QY	253	TETLSQVNPFRMFWTN-KASSNEPSLKELEKDP-VWKKLNQVNRVDILDRDLWARSR	310
Db	244	NEELANINPKVMILATDGTDKNRTKF----IDPAVWKS LKAVKDNKYVDVDRNKWLKSR	299
QY	311	GLISSEEMAKELVELSKK 328	
Db	300	GIIASESMAEDLEKIAEK 317	
RESULT 10			

```
US-10-282-122A-71291
; Sequence 71291, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haeelbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 71291
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Staphylococcus haemolyticus
US-10-282-122A-71291

Query Match      38.1%; Score 644; DB 15; Length 175;
Best Local Similarity 70.9%; Pred. No. 4e-36;
Matches 124; Conservative 23; Mismatches 24; Indels 4; Gaps 1;

Qy      5 VRGKILSVIGLLFVLIAATAACGNSSNS-----SKSSKDGVEIKHEEGTTKVPKPKR 60
Db      1 MKGFAGVALLFALVLVITACGNVNNNGSDSGNSKSSKDSIKIKHELGTTKVPKQAKR 60

Qy      61 VVLYEYSPDVALVDKVPVGIADNKNRIIKPLRDKIGKYSVGTTRKQNLSEISKJK 120
Db      61 VVALEFSPDALAALNVKVPVGVADNKNRIIKPLKEKIGKYSVGARKQNLSEISKJK 120

Qy      121 PDLIADNRRHKGIIYKDLNKIAPTIELKSPDGYNNENIDAPKTIKALGKEERGK 175
Db      121 PDLIADNRRHKGIIYKELSKIAPTIELKSPDGYNDNDIDAPKTIKALNKDDVQG 175

RESULT 11
US-10-282-122A-66890
; Sequence 66890, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
```

```
; APPLICANT: Malone, Cheryl
; APPLICANT: Haeelbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66890
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-10-282-122A-66890

Query Match      30.6%; Score 516.5; DB 15; Length 298;
Best Local Similarity 38.1%; Pred. No. 3.6e-27;
Matches 111; Conservative 65; Mismatches 98; Indels 17; Gaps 4;

Qy      43 VEIKHEEGTTKVPKPKRVVVLVLEYSFYDALVALDVKPVGIADNKNRIIKPLRDKIGKY 102
Db      23 VTVKQKGEFTLDSVPKRVVVALEYVYDALAQIGVSPGVADDDNDKTRILQKVRDKVQPW 82

Qy      103 TSVGTRKQPNLEIEISKJLKPDLIIADNRRHKGIIYKDLNKIAPTIELKSPDGYNNENIDAFK 162
Db      83 ESVGTRSQPSLEAISALKPDLIIADNRRHSAVYELKKIAPTIVVFNSRSHENYQENLETAQ 142

Qy      163 TISKALGKEERGKRLSEHDKKIEYKEITMDKNQVLPVAAKSGLLAHPNSNSYVQOF 222
Db      143 KIGOLLGKREMQARIATKHQODIADIAK--TLPKGKKAIGVSRQTNLNSRSHYAGGL 200

Qy      223 LSQLGFK--EALSDDVTTKGLSKYLKGPVLOQNTETLSQVNPFRMFMITNKASSNEPSLKE 280
Db      201 VEVLYGQMPKARADN-----QPNASVGLGQVAAEKPDMLLIHYR---DESIARK 247

Qy      281 LEKDPVWKULNAVQNQRVDILDRDLWARSGLISSEMAKVELVELSKDSK 331
Db      248 WENEALWKIIPAVKNGQVILANDNLWARAGIDAAEVMAKEVQDFVTKSAK 298

RESULT 12
US-10-282-122A-56054
; Sequence 56054, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
```



```

; APPLICANT: Wang, Liangu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56054
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Enterobacter cloacae
US-10-282-122A-56054

Query Match      27.2%; Score 460; DB 15; Length 300;
Best Local Similarity 35.8%; Pred. No. 2.5e-23;
Matches 101; Conservative 60; Mismatches 105; Indels 16; Gaps 4;

QY 45 IKHEGTTKPKPKRVVLEYSFVDALVDPKVGADNNKRIIKPLRDKIGKYS 104
DB 24 VQDEHGTFTLDTKPTQRIIVVLEISFADALAAVDVPIGADNDKRAKILPEVRAHLKPWQS 83

QY 105 VGTROPNLEISKLPDLIIADNNRHKGIVKDLNKIAPTIELKSPDGDYNNIDAFKTI 164
DB 84 VGTQAQPSLEATAALPKDLIIADSSRHAGIYALQOIAPVLLKSRNETYAENLHSAII 143

QY 165 SKALGKEEGKKRLEHDKKIEYKKEITMDKNQKVLPAVAKSGLLAHPNSYVQGFUS 224
DB 144 GEVVGKKREMQARLEQHKERMAQWSSQ--LPTGTRVAFGTSREQLFNLTQETWTGSLVLA 201

QY 225 QLGFKREALSDVTKGLSKYLKGPYL-QMNTETLSQVNPERRMFTMTNKASSNEPSLKELEK 283
DB 202 SLGL-----NVPAAAGASMPISGLEQLLAVNPALWLVAHYR---EESIVKRWQ 248

QY 284 DPVWKKLNKVNQRVDILDRDLWARSGLISSEEMAKELVEL 325
DB 249 DPLWQILTAQAQKQVAVSDSNARWARGIFAAERIAADTVKI 290

RESULT 13
US-09-815-242-10469
; Sequence 10469, Application US/09815242

; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10469
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10469

Query Match      27.2%; Score 460; DB 9; Length 302;
Best Local Similarity 35.5%; Pred. No. 2.5e-23;
Matches 100; Conservative 62; Mismatches 104; Indels 16; Gaps 4;

QY 45 IKHEGTTKPKPKRVVLEYSFVDALVDPKVGADNNKRIIKPLRDKIGKYS 104
DB 26 VQDEHGTFTLDTKPTQRIIVVLEISFADALAAVDVPIGADNDKRAKILPEVRAHLKPWQS 85

QY 105 VGTROPNLEISKLPDLIIADNNRHKGIVKDLNKIAPTIELKSPDGDYNNIDAFKTI 164
DB 86 VGTQAQPSLEATAALPKDLIIADSSRHAGIYALQOIAPVLLKSRNETYAENLHSAII 145

QY 165 SKALGKEEGKKRLEHDKKIEYKKEITMDKNQKVLPAVAKSGLLAHPNSYVQGFUS 224
DB 146 GEVVGKKREMQARLEQHKERMAQWSSQ--LPGTRVAFGTSREQLFNLTQETWTGSLVLA 203

QY 225 QLGFKREALSDVTKGLSKYLKGPYL-QMNTETLSQVNPERRMFTMTNKASSNEPSLKELEK 283
DB 204 SLGL-----NVPAAAGASMPISGLEQLLAVNPALWLVAHYR---EESIVKRWQ 250

QY 284 DPVWKKLNKVNQRVDILDRDLWARSGLISSEEMAKELVEL 325
DB 251 DPLWQILTAQAQKQVAVSDSNARWARGIFAAERIAADTVKI 292

RESULT 14
US-10-282-122A-56805
; Sequence 56805, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
```

```
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56805
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-282-122A-56805

Query Match      27.2%; Score 460; DB 15; Length 302;
Best Local Similarity 35.5%; Pred. No. 2.5e-23;
Matches 100; Conservative 62; Mismatches 104; Indels 16; Gaps 4;

Qy 45 IKHEGTTKPKPKRVVVLVSFVDALVDKPVGIADDNKKNRIKPLRDKIGKYS 104
Db 26 VQDEHGTFTLEKTPQRIIVLSLFSADALAAVDVIPIGIADNDNAKRILLPEVRAHLKPWQS 85
Qy 105 VGTRKQPNLEISIKLKPDLIIADNNRHKGIVKDLNKNKIAPTIELKSFQDGYNNENIDAFKTI 164
Db 86 VGTRAQPSLEIAALKPDLIIADSSRHAGVIALQOIAPVLLKSRNETYAENLQSAAIL 145
Qy 165 SKALGKEEGKRLEEHDKIEYEKKEITMDKNQKVLPAVAASKGLLAHPNSVYVGFLS 224
Db 146 GEMVQKKREMQARLEQHKERMAQASQ--LPKGTTRVAFGTSREQQFNLTQETWTGSLA 203
Qy 225 QLGPFKALSDDVTKGLSKYLKGPYL-QMNTETLSQVNPFRMFIMTNKASSNEPSLKELEK 283
Db 204 SLGL-----NVPAAMAGASMPSTGLQLLAVNPALLVAHYR---BESIVKRWQQ 250
Qy 284 DPVWKKLNAVKNQRIIDRLDLWARSRLGISSEEMAKSELVEL 325
Db 251 DPLWQMLTAQKQVQVAVSDSNWTWARMRGIPAAERIAADTVKI 292

RESULT 15
US-10-282-122A-69420
; Sequence 69420, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
```

```
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69420
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Pseudomonas syringae
US-10-282-122A-69420

Query Match      26.9%; Score 454; DB 15; Length 297;
Best Local Similarity 34.3%; Pred. No. 6.3e-23;
Matches 109; Conservative 60; Mismatches 109; Indels 40; Gaps 6;

Qy 17 LFVLIATAACGNSSNSSKSSKDGVEIKHEBGTQV--PKHPKRVVVLVSFVDALVA 74
Db 5 LLTLTLAGAA-----QAAPIDIDDDGQHKVHLPDTPKRVVVLVSFVDGLAS 49
Qy 75 LDVKEVGIADDNKKNRIKPLRDKIGKYSVGTRKQPNLEISIKLKPDLIIADNNRHKGI 134
Db 50 VGVTFVGAADPGDASRVLPKVRKAVGEMQSVGLRSQPNIEVIALKPLDIIADLGRHQA 109
Qy 135 YKDLNKIAPTIELKSFQDGYNNENIDAPKTIKALGKEEGKRLEEHDKIEYEKKEITM 194
Db 110 YNDLASIAPTILMSRGEDYQGSLSKSLAGLGMALGKGPQMQRARAEHQHKTVAEQIPA 169
Qy 195 DKNQKVLPAVAASKGLLAHPNSVYVGFPSQLGPFKALSDDVTKGLSKYLKGPYLQMTE 254
Db 170 DSN--VLFPGVAREDSFVHGPHSYAGSVLQIG-----LQVPEVRNNA 211
Qy 255 TLSQVNPFRMFIMTNK---ASSNEPSLKELEKDPVWKKLNAVKNQRIIDRLDLWARS 309
Db 212 PTEFVSLEQLLALDPNMLLVGHYRRPSIVDTWSKQPLMQVLGAVRNKQVAVEDGDSWARN 271
Qy 310 RGLISSEEMAKSELVELSK 327
Db 272 RGIWASEQIADDAALAVLK 289

RESULT 16
US-08-781-986A-5193
; Sequence 5193, Application US/08781986A
```



```

/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
/ COMPUTER: HP Vectra 486/33
/ OPERATING SYSTEM: MSDOS version 6.2
/ SOFTWARE: ASCII Text
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/10/329,624
/ FILING DATE: 27-Dec-2002
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/956,171
/ FILING DATE: October 20, 1997
/ APPLICATION NUMBER: 60/009,861
/ FILING DATE: January 5, 1996
/ APPLICATION NUMBER: 08/781,986
/ FILING DATE: January 3, 1997
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Mark J. Hyman
/ REGISTRATION NUMBER: 46,789
/ REFERENCE/DOCKET NUMBER: PB248P1D1
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (240) 314-1224
/ TELEFAX: (301) 309-8439
/ INFORMATION FOR SEQ ID NO: 5193:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 330 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 5193:

```

US-10-329-624-5193

```

Query Match      24.0%; Score 405; DB 15; Length 330;
Best Local Similarity 32.8%; Pred. No. 1.5e-19;
Matches 113; Conservative 65; Mismatches 132; Indels 34; Gaps 12;

QY  8 LKILSVIGLLFVLIIATAACGNSSNSSSKSSKDGVEIKHEEGTTKVPKPKRVVVLEYS 67
DB  5 IKML-VVTFLAVLV-LAGCSGNSNKQSSDNKDKEKTTSIKHAMGTTTEIKGPKRVVTLVYQG 62

QY  68 FVDALVALDVKPGVIADDNKNRIIKPLRDKIGKIVTSVTRKQPNLEISIKLKPDLIIAD 127
DB  63 ATDVAVSLGVKPVGAVESWTQKPKFEYIKNDLKDTKIVGQBPAPNLEISIKLKPDLIIAS 122

QY  128 NNRHGIYKDLNKIAPTIELKSPFGDYNENIDAFKTIKALGKEEGKKRLEEHDKKIEE 187
DB  123 KVRNEKVVYDQLSKIAPTIV---STDVTFKFK-DTTKLMGKALGKEKEAEDLLKKYDDKVA 178

QY  188 YKKEITWMDKNQKVLPAVAKSGLLAHPN-----SYVGQFLSOLGPKKEALSDDVTKGL 240
DB  179 FQK-----DAKAKYKDAWPLKASVNVFRADHTRIYAGGYAGEILNDLGFKR-----NKDL 228

QY  241 SKYLAG--PYLQMT--ETLSQVNPFRMFTMTNKASSNEPSL-----KELEKDPVWKKLNA 292
DB  229 QKQVDNGKDIIQLTSKESIIPLNADHIFVVKSDPNAKDAALVKKTESEWTSKKEWKNLDA 288

QY  293 VKNQRV-DILDRDLWARSGRLISSEEMAKELVELSKDKSKDNK 335
DB  289 VKNQVSDDLDEITWNLAGGYKSLKLIDDLVE--KLNIEKQSK 330

```

RESULT 19

```

US-10-470-048B-346
Sequence 346, Application US/10470048B
Publication No. US20050037444A1
GENERAL INFORMATION:
/ APPLICANT: MEINKE ET AL.
/ TITLE OF INVENTION: A METHOD FOR IDENTIFICATION, ISOLATION AND PRODUCTION OF
/ FILE REFERENCE: SONN:035US
/ CURRENT APPLICATION NUMBER: US/10/470,048B
/ CURRENT FILING DATE: 2003-07-25
/ NUMBER OF SEQ ID NOS: 603

```

```

/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 346
/ LENGTH: 330
/ TYPE: PRT
/ ORGANISM: Staphylococcus aureus
US-10-470-048B-346

```

```

Query Match      24.0%; Score 405; DB 17; Length 330;
Best Local Similarity 32.8%; Pred. No. 1.5e-19;
Matches 113; Conservative 65; Mismatches 132; Indels 34; Gaps 12;

QY  8 LKILSVIGLLFVLIIATAACGNSSNSSSKSSKDGVEIKHEEGTTKVPKPKRVVVLEYS 67
DB  5 IKML-VVTFLAVLV-LAGCSGNSNKQSSDNKDKEKTTSIKHAMGTTTEIKGPKRVVTLVYQG 62

QY  68 FVDALVALDVKPGVIADDNKNRIIKPLRDKIGKIVTSVTRKQPNLEISIKLKPDLIIAD 127
DB  63 ATDVAVSLGVKPVGAVESWTQKPKFEYIKNDLKDTKIVGQBPAPNLEISIKLKPDLIIAS 122

QY  128 NNRHGIYKDLNKIAPTIELKSPFGDYNENIDAFKTIKALGKEEGKKRLEEHDKKIEE 187
DB  123 KVRNEKVVYDQLSKIAPTIV---STDVTFKFK-DTTKLMGKALGKEKEAEDLLKKYDDKVA 178

QY  188 YKKEITWMDKNQKVLPAVAKSGLLAHPN-----SYVGQFLSOLGPKKEALSDDVTKGL 240
DB  179 FQK-----DAKAKYKDAWPLKASVNVFRADHTRIYAGGYAGEILNDLGFKR-----NKDL 228

QY  241 SKYLAG--PYLQMT--ETLSQVNPFRMFTMTNKASSNEPSL-----KELEKDPVWKKLNA 292
DB  229 QKQVDNGKDIIQLTSKESIIPLNADHIFVVKSDPNAKDAALVKKTESEWTSKKEWKNLDA 288

QY  293 VKNQRV-DILDRDLWARSGRLISSEEMAKELVELSKDKSKDNK 335
DB  289 VKNQVSDDLDEITWNLAGGYKSLKLIDDLVE--KLNIEKQSK 330

```

RESULT 20

```

US-10-967-189-16
Sequence 16, Application US/10967189
Publication No. US20050053995A1
GENERAL INFORMATION:
/ APPLICANT: Simpson et al.
/ TITLE OF INVENTION: Staphylococcus aureus Genes and Polypeptides
/ FILE REFERENCE: PB461USD2
/ CURRENT APPLICATION NUMBER: US/10/967,189
/ CURRENT FILING DATE: 2004-10-19
/ PRIOR APPLICATION NUMBER: US 10/278,946
/ PRIOR FILING DATE: 2002-10-24
/ PRIOR APPLICATION NUMBER: US 09/830,217
/ PRIOR FILING DATE: 2002-01-15
/ PRIOR APPLICATION NUMBER: PCT/US99/06199
/ PRIOR FILING DATE: 1999-03-18
/ PRIOR APPLICATION NUMBER: US 60/078,682
/ PRIOR FILING DATE: 1998-03-20
/ PRIOR APPLICATION NUMBER: US 60/080,296
/ PRIOR FILING DATE: 1998-04-01
/ PRIOR APPLICATION NUMBER: US 60/084,674
/ PRIOR FILING DATE: 1998-05-07
/ NUMBER OF SEQ ID NOS: 22
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 16
/ LENGTH: 330
/ TYPE: PRT
/ ORGANISM: Staphylococcus aureus
US-10-967-189-16

```

Query Match 24.0%; Score 405; DB 17; Length 330;

```

Best Local Similarity 32.8%; Pred. No. 1.5e-19;
Matches 113; Conservative 65; Mismatches 132; Indels 34; Gaps 12;

QY  8 LKILSVIGLLFVLIIATAACGNSSNSSSKSSKDGVEIKHEEGTTKVPKPKRVVVLEYS 67
DB  5 IKML-VVTFLAVLV-LAGCSGNSNKQSSDNKDKEKTTSIKHAMGTTTEIKGPKRVVTLVYQG 62

```

QY 68 FVDALVALDVKPGIADNDKNNRIIKPLRDKIGKYSVGTQKPNLEESIKLKPDLIAD 127
DB 63 ATDVAVSLGVKPGVAVESWTQPKFEYIKNDLKDQTKIVQBPAPNLEESIKLKPDLIVAS 122
QY 128 NNRHGIYKDLNKIAPTIELKSFDDGYNENIDAFKTIKALGKEBEGKRLBEHDKIIE 187
DB 123 KYRNEKVDOLSKIAPTV---STDIVPKFK-DTTLKMGKALGKEAEEDLLKKYDDKVA 178
QY 188 YKKEITMDKQKVLPAVAASKSLLAHPSN-----SYVGQFLSOLGPKFKEALSDDVTGKL 240
DB 179 FQK----DAKAKYKDAWPLKASVFNFRADHTRIYAGGYAGETLNDLGFKR-----NKDL 228
QY 241 SKVLKG--PYLQWNT-ETLSQVNPERRMTNKAASNEPSL-----KELEKDPVVKCLNA 292
DB 229 QKQVNGKDIILQTSKESILPMNADHIFVVKSDPNKADALVKKTESEWTSSTKEWNLDA 288
QY 293 VKNQV-DILDRDLWARSGLISSEEMAKELVELSKDSKDK 335
DB 289 VKNQVSDDLDEITWNLGGYKSLKLIDDLVE--KLNTEKQSK 330

RESULT 21
US-10-282-122A-45599
; Sequence 45599, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangau
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 45599
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-10-282-122A-45599

Best Local Similarity 34.4%; Pred. No. 1.4e-18;
Matches 116; Conservative 59; Mismatches 122; Indels 40; Gaps 13;
QY 9 KILSVIGLLFVLIIATAACGNSSSSSKESKD---GVEIKHEGTTKVPKPKRVVVL 64
DB 3 KILSI-----FIVWFLFVAGCCQOQKEKETKADKNQAITIKHAEGETKLDKPAKVVVL 58
QY 65 EYSFVDALVALDVKPGIADNDKNNRII---KPLRDKIGKYSVGTQKPNLEESIKLK 120
DB 59 EWWYSEDLLALGVQVPGWADIKYNNKWNTKPKSDVV---DVGTRQPNLEESIRLK 114
QY 121 PDLIIADNNRUKGIYKDLNKIAPTIEL---KSFDDGYNENIDAFKTIKALGKEBEGKRL 177
DB 115 PDLIIITASFRGKAIKNELEQIAPTVMFDPSTSNNDHFAEMTETFKQIAKAVKGBEGKVK 174
QY 178 LEEDHKKIEEYKKEI---TMDKQKVLPAVAAS---GLLAHPSNSYVGQFLSOLGFK 229
DB 175 LADMDKAPADAKAKIEKADLKDKKNIAMAQATKAKNPTFRILT--DNSLALQVTKLGG-- 230
QY 230 EALSDDVTGKLSKYLKGPYLOMTTETLSQVNPERR-MFIMTNKASNEPSLKELEKDPVWK 288
DB 231 --LNTTPEAGKSE--PDGFKQTTVESLQSVDSNFIYIVADEDNIFDTQLK---GNPAWE 283
QY 289 KLNQVQNRVDILDRDLWARSGLISSEEMAKELVEL 325
DB 284 ELKFKKENKMYKLGDTWIFG-GPESATSLATQVADV 319

RESULT 22
US-09-738-626-3875
; Sequence 3875, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENO, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 3875
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-3875

Query Match 19.6%; Score 331; DB 9; Length 315;
Best Local Similarity 28.6%; Pred. No. 1.6e-14;
Matches 90; Conservative 65; Mismatches 140; Indels 20; Gaps 7;
QY 13 VIGLLFVLIIATAACGNSSSSNS-SKESKDGVEIKHEGTTKVPKPKRVVLEYSFVDA 71
DB 15 VAPLLGVALLSSCSSTSSDESIGQPEVASTGYSVHAMGTETPTPTTPTTPTTPTTPTT 74
QY 72 LVALDVKPGIADNDKNNRIIKPLRDKIGKYSVGTQKPNLEESIKLKPDLIADNNRH 131
DB 75 LLALGITPVGATESGSENGFPAYLDELKDTESVGLTSEPNLEKIAALDPDLIIAGKVRH 134


```
Query Match      18.5%; Score 313; DB 15; Length 314;
Best Local Similarity 29.9%; Pred. No. 2.6e-13;
Matches 99; Conservative 65; Mismatches 135; Indels 32; Gaps 14;

QY 9 KILSVIGLFLVLIATAACGNSSNSSKSSKDGVEIKHEEGTTKVPKHPKRVVLEYSF 68
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 4 KLFILFTMLVLSVVGSSSQEESKAKEQPKTKV-VKHAKGEATIPVNPRIIVDLSGS- 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 69 VDALVALDVKPGVIADNNKRIIKPLDKIKGYTSVGTROP--NLREISKLPDLIIA 126
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 62 TEELLGLGHKPGVGT-A-NTYKDKIQKHLTEKLDGVKAVGWYAPKVDLEAVTALPKDLIL 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 127 DNNRHGKIYKOLNIAPIELKSPDQYNENIDAFKTIKALGKEEGKGLLEHDKKIE 186
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 -NNRQLKIYDQLEKAVPTVLTNLEDMR---GKFEVGLFDEKKADKIADYDKKAD 176
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 187 EYKKEITWDMKQKVLPAVAAGSGLLAHPSNSV-GQFLSQGLGFKALSDDV-----TKGLS 241
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 177 SLSKKIKKTKDDSEFMFAVT-----PQNFVYGSF---GYGDIIFNDLKLPAKGTG 226
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 242 KYLKGPIYQMTETLSQVNPFRMFTNKNASSNEPSLAKLEKDPVKKLNNAVKNQVRDIL 301
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 227 --LKQMAQVSLGLVAFQPDQMFIV-NFGGEADKVYEDYKNSAVKONKAVKNNHVYEV 283
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 302 DRDLW-ARSRGLISS-----EEMAKELVELSK 327
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 284 SNEVNTKAFNFIGDKMLDEIAKELILAKNK 314
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
RESULT 25
US-10-282-122A-45385
; Sequence 45385, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Lianguo
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 45385
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-10-282-122A-45385

Query Match      17.8%; Score 301; DB 15; Length 314;
Best Local Similarity 27.3%; Pred. No. 1.7e-12;
Matches 93; Conservative 66; Mismatches 106; Indels 76; Gaps 16;

QY 16 LLFLVLIAT----AACGNSSSNS--SKSSKDGVEIKHEEGTTKVPKHPKRVVLEYSFVD 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 9 LAWLVLVTSVLFVAFACSNEEKADAKAETERTVQAKGEIKIPANPKKIADLSGS-TE 67
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 71 ALVALDVKPGVIADNN-----KNRIIKPL-----RDKIGKYTSVGTTRKQPNLEIS 117
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 68 ELLIFGKMPIITANTSQEKIDAHEIKLKDVKPVGSAGWDKI-----NIEAVA 115
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 118 KLPDLIIADNNRHGKIYKOLNIAPIELKSPDQYNENIDAFKTIKALGKEEGK 177
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 116 AAKPDLILV--NNRQEKIYDQLSKIAPTVMKLT---PLDWRPKFEVVGQIFGKSKETKBW 171
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 178 LEEDHDKKIEYKKEITWDMKQKVLPAVAAGS-----LLAHPSNSVVGQFLSQGLFKEA 231
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 172 FKQYD-----EKASKLHDKIVAKTGDAKFMKMAAYP--NAP--RVYGDYGYGSV 216
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 232 LSDDV-----TKGLSKYLKGPYLOMTETLSQVNPFRMFTNKNASSNEPSLAKLEKOPVW 287
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 217 IFNDLKLPAVKGTP--TDKPLVQVQKEALIDYNPDYLFVFTTGDGDSQR--LKEFQEEISIW 272
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 288 KCLNAVKNQVRVDILDRDLWARSRLGISSEMAKELVELSKK 328
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 273 KNNNAVKNHV-----FTIKNEELNKGYPFLGKE 301
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
RESULT 26
US-10-282-122A-46193
; Sequence 46193, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Lianguo
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
```

```
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46193
; LENGTH: 264
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-10-282-122A-46193

Query Match          17.8%; Score 300; DB 15; Length 264;
Best Local Similarity 33.6%; Pred. No. 1.6e-12;
Matches 93; Conservative 50; Mismatches 102; Indels 32; Gaps 11;

Qy 65 EYSFVDALVALDVPVGIADNKKNR11-----KPLRDKIGKIVTSVGRKQPNLEISKLK 120
Db 2 EWVYSEDLALGVQPVGVADIKYNNKWNVTKTSKDVV-----DVGTRQPNLEISKLK 57

Qy 121 PDLIADNRRHGIYKOLNKIAPTIEL---KSPGDYNNENIDAFKTSKALGKBEKGKR 177
Db 58 PDLIITASFRGKA1KNELEQIAPTVMFDPSTSNNDHFAEMTETFKQIAKAVGKEEGKV 117

Qy 178 LEEHDKKIEBYKKI-----TWCKNQKVLPAVAKS-----GLLAHPSNSYVQFSLQSGFK 229
Db 118 LADMKAPADAKAKIEKADLKDKK1AMAQAFTAKNVPTFRILT--DNSLALQVTKKLG-- 173

Qy 230 EALSDDVTKGLSKYLKGPYLQMTETLSQVNPFR-MFIMTNKASSNEPSLKELEKDPVVK 288
Db 174 --LINTPEAGKSE--PDGFKTTVESLQSVQDSNFIYVADEDNIDFTQLK--GHPAWE 226

Qy 289 KINAVKQNRVDILDRDLWARSRGLISSEEMAKELVEL 325
Db 227 ELKFKKENKMYNLKGDWTFIG-GPESATSLATQVADV 262

RESULT 27
US-10-724-972A-7181
; Sequence 7181, Application US/10724972A
; Publication No. US20040147734A1
; GENERAL INFORMATION:
; APPLICANT: Doucette-Stamm, Lynn
; APPLICANT: Bush, David
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PATH03-16
; CURRENT APPLICATION NUMBER: US/10/724,972A
; PRIOR FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: 09/450,969
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: 09/134,001
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 7544
; SEQ ID NO 7181
; LENGTH: 351
; TYPE: PRT
; ORGANISM: S.epidermidis
US-10-724-972A-7181

Query Match          17.5%; Score 296.5; DB 18; Length 351;
Best Local Similarity 28.4%; Pred. No. 4e-12;
Matches 104; Conservative 70; Mismatches 131; Indels 61; Gaps 16;

Qy 1 QVESVRGUKILSVIGLLFVLVIATAACGNSSNSSKES-SKQGVBIKHE----- 48
Db 1 GVAHMKTVLFIALLSLVLVLTACSSNNSSNNSTSKKNSDSKETVTIKNSFEASGKENNGS 60
Qy 49 -----BGTTKVPKHPKRVVLEYSFVDALVALDVKVPVIADNKK-----KNRIIKPLRD 97
```

```
Db 61 DKKKISNTVEPKPNKNAVLDYCALDVL-----KEIGVADKVKGLPKGENNQSLPKFLD 115
Qy 98 --KIGKYTSVGRKQPNLEISKLKPDLI-IADNRRHGIYKOLNKIAPTIELKSPDGDY 154
Db 116 EFKDDKYINTGNLKEVNFVSAKPDVIFISGRTAQKNLDEPKAKPAKV-VYVGTS 174
Qy 155 NENI--DAFK---TISKALGKEBEGKKLEBHDHKKIEYKKEITMDKNQKVLPAVAKSG 209
Db 175 DDNLIDKMKNTENLGIYDKEDKAKKINKDLORKISDMKDK-TKDFNKKVMYLLVNEGE 233
Qy 210 LLAH-PSNSYVQFSLQSGFKEALSDDVTKGLSKYLKGPYLQ-MNTETLSQVNP----- 261
Db 234 LSTFGPGGRGGLVDFDTLGGKPA-----DKKVSFSPHCQNTNNYINKQNPVILAM 285
Qy 262 ERFPMITNKASSNEPSLKELEKDPVVKLANAVKQNRVDILDRDLWARSRGLISSEEMAKE 321
Db 286 DRGSVVGKATTN-----QVLKKNKVIKNKAVKSNHIYELDPKLYWYFSSG--SSTTTIKQ 338
Qy 322 LVLSK 327
Db 339 IDELNE 344

RESULT 28
US-10-282-122A-68785
; Sequence 68785, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malonebeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68785
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-10-282-122A-68785
```



```
; ORGANISM: Enterococcus faecalis
US-09-815-242-10905

Query Match      16.5%; Score 279; DB 9; Length 317;
Best Local Similarity 30.1%; Pred. No. 5.4e-11;
Matches 100; Conservative 60; Mismatches 130; Indels 42; Gaps 13;

QY 11 LSVIGLLFVLITATAACGNN---SSNSKESKOGVEIKHEEGTTKPKHPRVVLVLEYS 67
Db 11 VSMGLML-----SACQTKTKTADSAITETAKTEVTVDYNGQLTVPKNQKVVVFDNG 66
QY 68 FVDALVALDV--KPGVIADNDKKNRIIKPLRDKIKGYTSVGRKQPNLEESKLPDLII 125
Db 67 SLDYMDALGVGRVVGAPTKN-----IPAYLKKYQKVESAGGIKEPDLKINQLKPDII 121
QY 136 ADNRRHKGIYKDLNKIAPTIEL-----KSPDGDYNNIDAFKTIISKALKEBEGKRLSEH 181
Db 122 I-SGRQDYQEQLEKAIAPTIIYLAVDAKNPWASTKQNIETLGTI-----FKEEVAKEKITGL 177
QY 182 DKKIEYKKEITMDKNQKVLPAVAAKGLLAHPSNSYVQFLSQGLGFEALSDD-----VT 237
Db 178 EKXIADVKKQABASAN-NALVVLVNEGQLSAYKGSRGFLIHDTGPKAA--DDKIEAST 234
QY 238 KGLSKYLKGPYLOMNTETLSQVNPERFMIM-TNKASSNEPSLKELEKDPVWKKLNAVKNQ 296
Db 235 HQQS-----VSVEYVLEKNPGILFVVDRTKAIGGDDSKDNVAANELIKQTDAGKND 285
QY 297 RVDILORDLWARSRLISSEEMAKELVELSKK 328
Db 286 KVIMLQPDVWYLSGGGLSEMHIL---MIEDVKK 314

RESULT 33
US-10-282-122A-46179
; Sequence 46179, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46179
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-10-282-122A-46179

Query Match      16.2%; Score 274.5; DB 15; Length 315;
Best Local Similarity 27.4%; Pred. No. 1.1e-10;
Matches 92; Conservative 70; Mismatches 129; Indels 45; Gaps 13;

QY 12 SVIGLLFVLITATAACGNNSSNSKESKOGVEIKHEEG-TTKVPKHPKRVVVLVLEYSFVD 70
Db 10 AILSIFILLIGCSAKGDEKASATKEGKEIEVTDLSGRKVTFDKVPESFATLSMGDMN 69
QY 71 ALVALDVKPGIADNDKKNRIIKPLRDKIKGYTSVGRKQPNLEESKLPDLIIADNRR 130
Db 70 IHALGGKIVGRPD-----AKITLP--EDIKIQVIGNAHQNFQIASKLPDVLIIANGNF 123
QY 131 HKGIYKDLNKIAPTIE-----LKSPDG--DYNENIDAFKTIISKALKEBEGKRLSEH 181
Db 124 QKN-----PIVEGQGTQKVISSANSVQDIQKNIELYGTIMK---KEDKANELNQKI 172
QY 182 DKKIEYKKEITMDKNQKVLPAVAAKGLLAHPSNSYVQFLSQGLGFEALSDDVTGKLS 241
Db 173 NVQMKYKYEK---SDVKALLVYGAPGYLAALPTSLGSDILEKTGGKNIAD--FPEMK 226
QY 242 KYLGPYLOMNTETLSQVNPERFMIMTNKASSNEPSLK-----ELEKDPVWKKLNAVKNQ 296
Db 227 EYPQ--YAQLSVERIIEANPDVIVLIITH---GDPNSVKKAFEGEMMKNEAKNLDVAKON 281
QY 297 RVDILORDLWARSRLISSEEMAKELVELSKKDKK 332
Db 282 RVILPPDLFGSNPGTKVTEAM--DFMYKSIQDVKK 315

RESULT 34
US-09-071-035-40
; Sequence 40, Application US/09071035
; Publication No. US20020045737A1
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 40:
```

```
;
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 296 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-09-071-035-40

Query Match      15.9%; Score 268; DB 9; Length 296;
Best Local Similarity 30.3%; Pred. No. 2.8e-10;
Matches 96; Conservative 54; Mismatches 129; Indels 38; Gaps 12;

Qy 26 CGNN---SSNSSKSSKSGVEIKHEEGTTKVPKPKRVVVVLEYSFVDALVALDV--KPV 80
Db 1 CQTNKTADSATTETAKTEVTVDKTNQGLTVPKPKRVVVVFNGLSMDTMDALGVGDRV 60

Qy 81 GIADDNKNRIILKPLDKIGKYSVGTGRKQPNLEISKLKPDLLIADNNRHKGIYKDLNK 140
Db 61 GAPTGN-----IPAYLKYYQKVESAGGIKEPDLEKINQKLPDLLII-SGRQDDYQEQQLKA 114

Qy 141 IAPTIEL-----KSPDGDYNNENIDAFKTIISKALGKEEKGKKLEHDKKIEYKKEITMDK 196
Db 115 IAPTIYLAVDKPNWASTKQNIETLGTI---FDKEEVAKEKITGLEKEIADVKKQAEASA 171

Qy 197 NQKVLPAVAKSGLLAHPNSNVVGFSLQGLFKEALSDD----VTKGLSKYLKGPYLOMN 252
Db 172 N-NALVVLVNEGQLSAYGKGRFGLIHDTFGFKAA--DDKTEASTHGQS-----VS 219

Qy 253 TETLSQVNPFRMFIM-TNKASSNPSPSLKELEKDPVWKKLNVAQNORVDILDRDLWARSRG 311
Db 220 YEYVLEKNPGLFVVDRTKAIAGDDSDKNVAANELIOKTDAGKNDKVIQLQPDVWYLSGG 279

Qy 312 LISSEEMAKELVELSKK 328
Db 280 GLESMHL---MIEDVKK 293

RESULT 35
US-10-206-576-40
; Sequence 40, Application US/10206576
; Publication No. US20030017495A1
; GENERAL INFORMATION:
;   APPLICANT: Choi et al.
;   TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
;   NUMBER OF SEQUENCES: 497
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Human Genome Sciences, Inc.
;     STREET: 9410 Key West Avenue
;     CITY: Rockville
;     STATE: Maryland
;     COUNTRY: USA
;     ZIP: 20850
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: CD-R
;     COMPUTER: Dell Latitude
;     OPERATING SYSTEM: Windows 98
;     SOFTWARE: ASCII Text
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/10/206,576
;     FILING DATE: 29-Jul-2002
;     CLASSIFICATION: <Unknown>
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 09/071,035
;     FILING DATE: 1998-05-04
;     APPLICATION NUMBER: US 60/046,655
;     FILING DATE: 1997-05-16
;     APPLICATION NUMBER: US 60/044,031
;     FILING DATE: 1997-05-06
;     APPLICATION NUMBER: US 60/066,009
;     FILING DATE: 1997-11-14
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Hyman, Mark J.
;     REGISTRATION NUMBER: 46,789
```

```
;
; REFERENCE/DOCKET NUMBER: PB369P1D1
;
; INFORMATION FOR SEQ ID NO: 40:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 296 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;     MOLECULE TYPE: protein
;   SEQUENCE DESCRIPTION: SEQ ID NO: 40:
US-10-206-576-40

Query Match      15.9%; Score 268; DB 14; Length 296;
Best Local Similarity 30.3%; Pred. No. 2.8e-10;
Matches 96; Conservative 54; Mismatches 129; Indels 38; Gaps 12;

Qy 26 CGNN---SSNSSKSSKSGVEIKHEEGTTKVPKPKRVVVVLEYSFVDALVALDV--KPV 80
Db 1 CQTNKTADSATTETAKTEVTVDKTNQGLTVPKPKRVVVVFNGLSMDTMDALGVGDRV 60

Qy 81 GIADDNKNRIILKPLDKIGKYSVGTGRKQPNLEISKLKPDLLIADNNRHKGIYKDLNK 140
Db 61 GAPTGN-----IPAYLKYYQKVESAGGIKEPDLEKINQKLPDLLII-SGRQDDYQEQQLKA 114

Qy 141 IAPTIEL-----KSPDGDYNNENIDAFKTIISKALGKEEKGKKLEHDKKIEYKKEITMDK 196
Db 115 IAPTIYLAVDKPNWASTKQNIETLGTI---FDKEEVAKEKITGLEKEIADVKKQAEASA 171

Qy 197 NQKVLPAVAKSGLLAHPNSNVVGFSLQGLFKEALSDD----VTKGLSKYLKGPYLOMN 252
Db 172 N-NALVVLVNEGQLSAYGKGRFGLIHDTFGFKAA--DDKTEASTHGQS-----VS 219

Qy 253 TETLSQVNPFRMFIM-TNKASSNPSPSLKELEKDPVWKKLNVAQNORVDILDRDLWARSRG 311
Db 220 YEYVLEKNPGLFVVDRTKAIAGDDSDKNVAANELIOKTDAGKNDKVIQLQPDVWYLSGG 279

Qy 312 LISSEEMAKELVELSKK 328
Db 280 GLESMHL---MIEDVKK 293

RESULT 36
US-10-912-362-40
; Sequence 40, Application US/10912362
; Publication No. US20050043528A1
; GENERAL INFORMATION:
;   APPLICANT: Choi et al.
;   TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
;   NUMBER OF SEQUENCES: 497
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Human Genome Sciences, Inc.
;     STREET: 14200 Shady Grove Road
;     CITY: Rockville
;     STATE: Maryland
;     COUNTRY: USA
;     ZIP: 20850
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: CD-R
;     COMPUTER: Dell Latitude
;     OPERATING SYSTEM: Windows 2000
;     SOFTWARE: ASCII Text
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/10/912,362
;     FILING DATE: 06-Aug-2004
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 10/206,576
;     FILING DATE: 2002-07-29
;     APPLICATION NUMBER: US 09/071,035
;     FILING DATE: 1998-05-04
;     APPLICATION NUMBER: US 60/046,655
;     FILING DATE: 1997-05-16
;     APPLICATION NUMBER: US 60/044,031
;     FILING DATE: 1997-05-06
;     APPLICATION NUMBER: US 60/066,009
```

FILING DATE: 1997-11-14
ATTORNEY/AGENT INFORMATION:
NAME: Hyman, Mark J.
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB369P1D2
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 296 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 40:
US-10-912-362-40

Query Match 15.9%; Score 268; DB 17; Length 296;
Best Local Similarity 30.3%; Pred. No. 2.8e-10;
Matches 96; Conservative 54; Mismatches 129; Indels 38; Gaps 12;

QY 26 CGNN---SSNSKSSKDGVEIKHEEGTTKVPKPKRVVLEYSFVDALVALDV--KPV 80
DB 1 COTNKKTADSATETTTAKTEVTIKTNGQLTVPKPKRVVFDNGSLDTMDALGVGDRVV 60

QY 81 GIADNKNRIIKPLRDKIKGYTSVCTRKQPNLEESIKLKPDLIIADNNRHKGIIYKDLNK 140
DB 61 GAPTNN-----IPAYLKQYKVESAGSAGIKPEPLEKINQLKPDLIIII-SGRQDYQEQLKA 114

QY 141 IAPTIEL----KSGFDGYNENIDAFKTSKALGKEEGKRLKEHDKKIEEYKKEITMDK 196
DB 115 IAPTIYLAVDKPNWASTKQNIETLGTI---PDKEEVAKEKITGLEKEIADYVKQAEASA 171

QY 197 NQKVLPAVAKSGLLAHPNSYVGFSLQSGKEALSDDP-----VTKGLSKYLKGPYLONN 252
DB 172 N-NALVVLVNEGQLSAYGKSRFGLIHDTFGKAA--DDKIEASTHGQS-----VS 219

QY 253 TETLSQVNPFRMFI--TNKASSNEPSLKELEKDPVWKLNNAVKNQVRVDILDRDLWARSRG 311
DB 220 YEIVLEKNPILFVDRTRKAIAGDDSKDNVAANELIKQTDAGKNDKVMQLQPDVVVLSG 279

QY 312 LISSEEMAKELVELSKK 328
DB 280 GLESMHL---MIEDVKK 293

RESULT 37
US-09-738-626-4403
Sequence 4403, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAKO
APPLICANT: SENO, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OKAZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 4403
LENGTH: 306

TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-4403

Query Match 15.7%; Score 266; DB 9; Length 306;
Best Local Similarity 25.9%; Pred. No. 4e-10;
Matches 82; Conservative 60; Mismatches 135; Indels 40; Gaps 7;

QY 17 LFVLIATA---ACGNSSSSNSKSSKDGVEIKHEEGTTKVPKPKRVVLEYSFVDAL 72
DB 11 LIVLAAALPLTACSSSSSEEAATSSATR--EFTDAGHTTEVPENPQRVVLEPLELDTA 68

QY 73 VALDVKPVGIADNKNRIIKPLRDKIKGYTSVCTRKQPNLEESIKLKPDLIIADNNRHK 132
DB 69 IALGITPVGAANNVTGI--PAYLGVDTGIEFVTVSEPNIEAIAALEPDLILGTDSRHA 126

QY 133 GIYKDLANKIAPTIELKSPFDGYNENIDAFKTSKALGKEEGKRLKEHDKKIEEYKKEI 192
DB 127 EYDRLESIAPTVFMVTHVDPKQNV---VFIGDALGKKQSEDLIQGFNDKCEIKSEH 183

QY 193 TMDKNQKVLPAVAKSGLLAHPNSYVGFSLQSGKEALSDDVTGKLSKYLKGPYLONN 252
DB 184 DVEGKTVNMIRPRDEQTMSLYGPSTSPAGSSLEACAG-----LTIPDQEWK 227

QY 253 TETLSQVNPFRMFI-----MTNKASSNEPSLKELEKDPVWKLNNAVKNQVRVDILDRDL 305
DB 228 DDLQADIAPENFMLATADYVFVATDVTDENELPEVIRE-----NREQFPFSLTLVDTSY 281

QY 306 WARSGLISSEEMAKEL 322
DB 282 WVGVGVLGGSGKVLIEDI 298

RESULT 38
US-10-627-476-572
Sequence 572, Application US/10627476
Publication No. US20040030116A1
GENERAL INFORMATION:
APPLICANT: Pompeius, Mark
APPLICANT: Kroger, Burkhard
APPLICANT: Schoder, Hartwig
APPLICANT: Zelder, Oskar
APPLICANT: Haberer, Gregor
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
TITLE OF INVENTION: TRANSPORT
FILE REFERENCE: BGI-125CPCN
CURRENT APPLICATION NUMBER: US/10/627,476
CURRENT FILING DATE: 2003-07-25
PRIOR APPLICATION NUMBER: 09/602,787
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: USSN 60/141031
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: DE 19931454.3
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931478.0
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931563.9
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19932122.1
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932124.8
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932125.6
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932128.0
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932180.9
PRIOR FILING DATE: 1999-07-09
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 678
SEQ ID NO 572
LENGTH: 306

Tue Sep 6 09:37:22 2005

[illegible]

Search completed: August 25, 2005, 09:40:02
Job time : 166 secs

This Page Blank (usptol)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 09:18:46 ; Search time 167 Seconds
(without alignments)
775.837 Million cell updates/sec

Title: US-10-724-972A-6352

Perfect score: 1690

Sequence: 1 GVESVRLGKILSVIGLFLVL.....EEMAKELVELSKDSKKDNK 335

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 16Dec04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1690	100.0	335	8	AdS07057 Staphyloc
2	1688	98.7	331	6	Abu43214 Protein e
3	1242	73.5	337	4	Auu36661 Staphyloc
4	1233	73.0	337	6	Abu41865 Staphyloc
5	1232	72.9	327	6	Abu15904 Protein e
6	1228	72.7	337	6	Abu38894 Surface e
7	1214	71.8	348	6	Abm73426 Staphyloc
8	1181	69.9	309	4	Auu33963 Staphyloc
9	726	43.0	319	4	Auu33753 Staphyloc
10	726	43.0	319	4	Auu36529 Staphyloc
11	724.5	42.9	319	6	Abu19052 Pathogen
12	724.5	42.9	319	6	Abu16209 Protein e
13	724.5	42.9	325	6	Abu38895 Surface e
14	724.5	42.9	325	6	Abu41867 Staphyloc
15	724.5	42.9	325	6	Abm72855 Staphyloc
16	644	38.1	175	6	Abu43367 Protein e
17	516.5	30.6	298	6	Abu38966 Protein e
18	476	28.2	303	6	Abm68354 Photorhab
19	460	27.2	300	6	Abu28130 Protein e
20	460	27.2	302	4	Auu34876 E. coli c
21	460	27.2	302	6	Abu28881 Protein e
22	454	26.9	297	6	Abu41496 Protein e
23	405	24.0	330	2	Aau89745 Staphyloc
24	405	24.0	330	2	Aay42331 Staphyloc
25	405	24.0	330	2	Aay1824 Staphyloc

26	405	24.0	330	6	ABJ19069	Abj19069 Pathogen
27	405	24.0	330	6	ABR38893	AbR38893 Surface e
28	405	24.0	330	6	ABR41864	AbR41864 Staphyloc
29	405	24.0	330	7	ADF43556	Adf43556 Staphyloc
30	391	23.1	324	6	ABU17675	Abu17675 Protein e
31	375	22.2	1132	4	ABG08851	Abg08851 Novel hum
32	350	20.7	932	4	ABG13148	Abg13148 Novel hum
33	343	20.3	287	6	ABM73050	Abm73050 Staphyloc
34	331	19.6	315	4	AAG90121	Agg90121 C glutami
35	327	19.3	321	6	ABU25343	Abu25343 Protein e
36	313	18.5	314	6	ABU18359	Abu18359 Protein e
37	307.5	18.2	352	7	ADC94889	Adc94889 E. faeciu
38	301	17.8	314	6	ABU17461	Abu17461 Protein e
39	300	17.8	264	6	ABU18269	Abu18269 Protein e
40	296.5	17.5	351	5	ABP40647	Abp40647 Staphyloc
41	296.5	17.5	351	8	ADS07886	AdS07886 Staphyloc
42	296	17.5	342	6	ABU40861	Abu40861 Protein e
43	296	17.5	351	7	ADF05577	Adf05577 Bacterial
44	292	17.3	347	4	AGS82087	AgS82087 S. epider
45	292	17.3	347	4	AAG81684	Agg81684 S. epider

ALIGNMENTS

RESULT 1

ADS07057
ID ADS07057 standard; protein; 335 AA.

XX ADS07057;

AC ADS07057;

DT 04-NOV-2004 (first entry)

XX Staphylococcus epidermis polypeptide seqid 6352.

XX antibacterial; vaccine; antisense therapy; Staphylococcus epidermidis;
KW recombinant expression vector; infection; computer readable medium;
KW computer based system.

XX Staphylococcus epidermidis.

XX US2004147734-A1.

XX 29-JUL-2004.

XX 01-DEC-2003; 2003US-00724972.

XX 08-NOV-1997; 97US-0064964P.

XX 13-AUG-1998; 98US-00134001.

XX 29-NOV-1999; 99US-00450969.

XX (DOUC/) DOUCETTE-STAMM L.

XX (BUSH/) BUSH D.

XX Doucette-Stamm L, Bush D;

XX WPI; 2004-580138/56.

XX N-PSDB; ADS03285.

XX New isolated polypeptide and encoding nucleic acid derived from
PT Staphylococcus epidermidis, useful for diagnosing, preventing and/or
PT treating an S. epidermidis bacterial infection.

XX Claim 17; SEQ ID NO 6352; 741pp; English.

XX The invention describes an isolated nucleic acid comprising a nucleotide
XX sequence with any of 3772 fully defined nucleotide sequences (SEQ ID NO:
XX 1-3772) and encoding an Staphylococcus epidermidis polypeptide with any
XX of 3772 fully defined amino acid sequences (SEQ ID NO: 3772-7544) as
XX given in the specification. Also described are: a recombinant expression
XX vector; a cell comprising a recombinant expression vector of (1);
XX producing an S. epidermidis polypeptide; an isolated nucleic acid
XX comprising a nucleotide sequence of at least 8 nucleotides in length; a

CC vaccine composition for prevention or treatment of an S. epidermidis
CC infection, comprising a nucleic acid cited above and a carrier; treating
CC a subject for S. epidermidis infection; a recombinant or substantially
CC pure preparation of an S. epidermidis polypeptide or its fragment; a
CC vaccine composition for prevention or treatment of an S. epidermidis
CC infection; detecting the presence of a Staphylococcus nucleic acid in a
CC sample; a computer readable medium having recorded in it the nucleotide
CC sequences with SEQ ID NO: 1-3772 or its fragments; a computer based
CC system for identifying fragments of the Staphylococcus genome of
CC commercial importance; a computer based system for identifying fragments
CC of the Staphylococcus plasmids of commercial importance; identifying
CC commercially important nucleic acid fragments of the Staphylococcus
CC genome and/or plasmids; and identifying an expression modulating fragment
CC of the Staphylococcus genome and/or plasmids. The methods and
CC compositions of the present invention are useful for the diagnosis,
CC prevention and/or treatment of an Staphylococcus epidermidis bacterial
CC infection. This is the amino acid sequence of a S. epidermis protein of
CC the invention.

XX Sequence 335 AA;

Query Match 100.0%; Score 1690; DB 8; Length 335;
Best Local Similarity 100.0%; Pred. No. 1.5e-129;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GVESVRGLKILSVIGLLFVLITATAACGNNSSNSSKSSKDGVEIKHEGTTKVPKPKR 60
DB 1 GVESVRGLKILSVIGLLFVLITATAACGNNSSNSSKSSKDGVEIKHEGTTKVPKPKR 60
QY 61 VVLEYSFVDALVALDVPVGIADNKKRIIKPLRDKIGKTVSGTRKQPNLEISK 120
DB 61 VVLEYSFVDALVALDVPVGIADNKKRIIKPLRDKIGKTVSGTRKQPNLEISK 120
QY 121 PDLIADNNRHKGIIYKDLNKTAPTIELKSFQDYNENIDAFKTIKALGKEEGKRL 180
DB 121 PDLIADNNRHKGIIYKDLNKTAPTIELKSFQDYNENIDAFKTIKALGKEEGKRL 180
QY 181 HDKKEEYKEITMDKNQKVLPAVAAGSGLLAHPNSVYVQFLSQGFKEALSDVT 240
DB 181 HDKKEEYKEITMDKNQKVLPAVAAGSGLLAHPNSVYVQFLSQGFKEALSDVT 240
QY 241 SKYLKGPYLQWNTETLSQVNPFRMFTMTNKASSNEPSLKELEKDPVWKKLN 300
DB 241 SKYLKGPYLQWNTETLSQVNPFRMFTMTNKASSNEPSLKELEKDPVWKKLN 300
QY 301 LDRDLWARSRLISSEEMAKELVELSKDKDNK 335
DB 301 LDRDLWARSRLISSEEMAKELVELSKDKDNK 335

RESULT 2
ABU43214
ID ABU43214 standard; protein; 331 AA.

XX AC ABU43214;

XX DT 19-JUN-2003 (first entry)

XX DE Protein encoded by Prokaryotic essential gene #28741.

XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS Staphylococcus epidermidis.

XX PN WO20027183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-0009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX PA (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.
XX N-PSDB; ACA47084.

XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 71138; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid;
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than S. aureus, S. typhimurium,
XX K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
XX the target prokaryotic essential genes. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

XX Sequence 331 AA;

Query Match 98.7%; Score 1668; DB 6; Length 331;
Best Local Similarity 99.7%; Pred. No. 9.2e-128;
Matches 330; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 VRGLKILSVIGLLFVLITATAACGNNSSNSSKSSKDGVEIKHEGTTKVPKPKR 64
DB 1 MRGLKILSVIGLLFVLITATAACGNNSSNSSKSSKDGVEIKHEGTTKVPKPKR 60

QY 65 BVSFVDALVALDVPVGIADNKKRIIKPLRDKIGKTVSGTRKQPNLEISK 124

DB 61 BVSFVDALVALDVPVGIADNKKRIIKPLRDKIGKTVSGTRKQPNLEISK 120

QY 125 IADNNRHKGIIYKDLNKTAPTIELKSFQDYNENIDAFKTIKALGKEEGKRL 184

DB 121 IADNNRHKGIIYKDLNKTAPTIELKSFQDYNENIDAFKTIKALGKEEGKRL 180

QY 185 IEEYKKEITMDKNQKVLPAVAAGSGLLAHPNSVYVQFLSQGFKEALSDVT 244

DB 181 IEEYKKEITMDKNQKVLPAVAAGSGLLAHPNSVYVQFLSQGFKEALSDVT 240

QY 245 KGPYLQWNTETLSQVNPFRMFTMTNKASSNEPSLKELEKDPVWKKLN 304


```

CC diagnostic assays
XX Sequence 337 AA;
SQ

Query Match 73.0%; Score 1233; DB 6; Length 337;
Best Local Similarity 71.6%; Pred. No. 3.3e-92;
Matches 235; Conservative 43; Mismatches 50; Indels 0; Gaps 0;

Qy 4 SVRGKILSVGLLVLTATACGNNSSSSKSSKDGVEIKHEEGTTKVPKPKRVVV 63
Dy :||||| :||| : :||| : :||| : :||| : :||| : :||| : :|||
10 TMRGLKTFSLGLVALLVAACGNTDSSKSSSTKDTISVKDENGTVKVPKDAKRIW 69
Qy 64 LEYSFVALVALDVKPGIADDDKKNRIIKPLDKIGKYSVGTGRKQNLSEISKLPDL 123
Dy :||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
70 LEYSFADALAALDVKPGIADDDKKNRIIKPLDKIGKYSVGTGRKQNLSEISKLPDL 129
Qy 124 IADNNRHKGIVKDLNKIAPTIELKSPDGYNNENIDAFKTSKALGKEEGKRLLEHDK 183
Dy :||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
130 IADSSRHKGINKELNKIAPTIELKSPDGYNNENIDAFKTSKALGKEEGKRLLEHDK 189
Qy 184 KIEYKKEITWDMKNQKVLPAVAAGSLLAHPSNSYVGQFLSQLPKEALSDVTKGLSKY 243
Dy :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
190 LINKYDEIKFDRNQKVLPAVAAGSLLAHPSNSYVGQFLNELGPKNALSDVTKGLSKY 249
Qy 244 LKGPVLOMNTETLSOVNPERFIMTNKASSNEPSLKELEKDPVWKKLNAVKNQVVDILDR 303
Dy :||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
250 LKGPVLOMNTETLSOVNPERFIMTNKASSNEPSLKELEKDPVWKKLNAVKNQVVDILDR 309
Qy 304 DLWARSRLISSEEMAKELVELSKKDSK 331
Dy :||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
310 DVWARSRLISSEEMAKELVELSKKDSK 337

RESULT 5
ABU15904
ID ABU15904 standard; protein; 327 AA.
XX
AC ABU15904;
XX
DT 19-JUN-2003 (first entry)
DE Protein encoded by Prokaryotic essential gene #1431.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
XX
PR 06-SEP-2001; 2001US-00948993.
XX
PR 25-OCT-2001; 2001US-0342923P.
XX
PR 08-FEB-2002; 2002US-00072851.
XX
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zvekind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR N-PSDB; ACA19774.
XX
WPI; 2003-029926/02.
XX
New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 43828; 1766pp; English.
XX

```

```

CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 327 AA;

```

```

Query Match 72.9%; Score 1232; DB 6; Length 327;
Best Local Similarity 71.9%; Pred. No. 3.8e-92;
Matches 235; Conservative 42; Mismatches 50; Indels 0; Gaps 0;

Qy 5 VRGKILSVGLLVLTATACGNNSSSSKSSKDGVEIKHEEGTTKVPKPKRVVV 64
Dy :||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1 MRGLKTFSLGLVALLVAACGNTDSSKSSSTKDTISVKDENGTVKVPKDAKRIW 60
Qy 65 BYSFVDALVALDVKPGIADDDKKNRIIKPLDKIGKYSVGTGRKQNLSEISKLPDL 124
Dy :||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
61 EYSFADALAALDVKPGIADDDKKNRIIKPLDKIGKYSVGTGRKQNLSEISKLPDL 120
Qy 125 IADNNRHKGIVKDLNKIAPTIELKSPDGYNNENIDAFKTSKALGKEEGKRLLEHDK 184
Dy :||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
121 IADSSRHKGINKELNKIAPTIELKSPDGYNNENIDAFKTSKALGKEEGKRLLEHDK 180
Qy 185 IEYKKEITWDMKNQKVLPAVAAGSLLAHPSNSYVGQFLSQLPKEALSDVTKGLSKYL 244
Dy :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
181 INKYDEIKFDRNQKVLPAVAAGSLLAHPSNSYVGQFLNELGPKNALSDVTKGLSKYL 240
Qy 245 KGPVLOMNTETLSOVNPERFIMTNKASSNEPSLKELEKDPVWKKLNAVKNQVVDILDR 304
Dy :||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
241 KGPVLOMNTETLSOVNPERFIMTNKASSNEPSLKELEKDPVWKKLNAVKNQVVDILDR 300
Qy 305 LWARSRLISSEEMAKELVELSKKDSK 331
Dy :||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
301 VWARSRLISSEEMAKELVELSKKDSK 327

RESULT 6
ABR38894
ID ABR38894 standard; protein; 337 AA.
XX
AC ABR38894;
XX
DT 10-MAY-2003 (first entry)
DE Surface exposed immunogenic polypeptide (SEIP) # SEQ ID 20.
XX

```

KW SEIP; surface exposed immunogenic polypeptide; iron-binding; receptor;
 KW antibacterial; fungicide; protozoacide; infection; disinfection.
 XX Unidentified.
 OS WO2003006672-A2.
 PN 23-JAN-2003.
 XX
 PD 10-JUL-2002; 2002WO-US021663.
 XX
 PF 10-JUL-2001; 2001US-0304390P.
 XX
 PR (DSQU-) D-SQUARED BIOTECHNOLOGIES INC.
 XX
 PA Scott DL;
 XX
 PI WPI; 2003-221756/21.
 XX
 DR Isolating microbial target surface exposed immunogenic polypeptides, by
 PT immunizing host with microbial membrane-associated polypeptides to
 PT produce antibodies and using antibodies to isolate target polypeptides.
 XX
 PS Claim 15; Page 40-41; 46pp; English.
 XX
 CC The invention relates to the isolation of target surface exposed
 CC immunogenic polypeptides (SEIP) from microorganisms. The SEIP's of the
 CC invention are receptors of iron-binding molecules. The activity of
 CC polypeptides of the invention may be described as antibacterial,
 CC fungicide and protozoacide. Pharmaceutical compositions based on the
 CC polypeptides of the invention are useful for treating infections and
 CC diseases caused by microbes including bacteria, fungi and protozoa in a
 CC patient. They are also useful for topical disinfection in an animal or
 CC mammal e.g. human. The current sequence represents an SEIP of the
 CC invention
 XX
 SQ Sequence 337 AA;
 Query Match 72.7%; Score 1228; DB 6; Length 337;
 Best Local Similarity 71.3%; Pred. No. 8.4e-92;
 Matches 234; Conservative 43; Mismatches 51; Indels 0; Gaps 0;
 QY 4 SVRGLKILSVIGLLFVLIATAACGNSSNSSKSSKDGVEIKHEGTTKVPKPKRVVV 63
 DB 10 TWRGLKTFILGLIIVALLVAACGNTDSSKSSSTKDTISVKDENGTVKYPKDAKRIIV 69
 QY 64 LEYSFVDALVALDVKPVGIIADNKKRIIIPKLRDKIGKTVTSVGTQKPNLEISIKLPDL 123
 DB 70 LEYSFADALVALDVKPVGIIADNKKRIIIPKVRKIGDVTSGTRKQPNLEISIKLPDL 129
 QY 124 IADNRRHGIYKDLNKIAPTIELKSPDGYNENIDAFKTIISKALGKEEGKKLEEDHK 183
 DB 130 IADSSRHGINKELNKIAPTIELKSPDGYNENIDAFKTIISKALNKEKEGKRLAEHDK 189
 QY 184 KIEEYKKEITMDKNQKVLPAVAKSGLLAHPNSYVVGQFLSOLGPKFKAALSDVTKGLSKY 243
 DB 190 LINKYKDEIKFDRNQKVLPAVAKAGLLAHPNYSYVVGQFLNGLFKNALSDVTKGLSKY 249
 QY 244 LKGPYLQMTTETLSQVNPFRMTNWKASNEPSLKELEKDPVWKKNVAVKQNRVILDR 303
 DB 250 LKGPYLQMTTETLSQVNPFRMTNWKASNEPSLKELEKDPVWKKNVAVKQNRVILDR 309
 QY 304 DLWARSRLISSEEMAKELVELSKKQSK 331
 DB 310 DVWARSRLISSEEMAKELVELSKKQSK 337
 RESULT 7
 ABM73426
 ID ABM73426 standard; protein; 348 AA.
 XX
 AC ABM73426;
 XX

DT 20-NOV-2003 (first entry)
 XX Staphylococcus aureus protein #2666.
 DE
 XX Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
 KW enzymatic assay; antibiotic target.
 XX
 OS Staphylococcus aureus.
 XX
 PN WO200294868-A2.
 XX
 PD 28-NOV-2002.
 XX
 PF 27-MAR-2002; 2002WO-IB002637.
 XX
 PR 27-MAR-2001; 2001GB-00007661.
 XX
 PA (CHIR-) CHIRON SPA.
 XX
 PI Masignani V, Mora M, Scarselli M;
 XX
 DR WPI; 2003-120786/11.
 DR N-PSDB; ACF74986.
 XX
 PT New Staphylococcus aureus protein, useful as a vaccine for treating or
 PT preventing Staphylococcal infection, specifically an infection caused by
 PT S. aureus, e.g. sepsis.
 XX
 PS Claim 1; SEQ ID NO 5332; 49pp; English.
 XX
 CC The invention relates to novel genes and encoded proteins from
 CC Staphylococcus aureus. A composition comprising the S. aureus protein, a
 CC nucleic acid encoding the protein, or an antibody to the protein, is
 CC useful as a pharmaceutical, particularly as a vaccine for treating or
 CC preventing infection due to Staphylococcus bacteria, specifically an
 CC infection caused by S. aureus. The composition is particularly useful for
 CC treating or preventing sepsis in a patient. The composition can also be
 CC used for diagnostics. The protein is also used in an assay for enzymatic
 CC studies and as a target for antibiotics. This sequence represents one of
 CC the novel S. aureus proteins of the invention
 XX
 SQ Sequence 348 AA;
 Query Match 71.8%; Score 1214; DB 6; Length 348;
 Best Local Similarity 71.8%; Pred. No. 1.2e-90;
 Matches 232; Conservative 41; Mismatches 50; Indels 0; Gaps 0;
 QY 9 KILSVIGLLFVLIATAACGNSSNSSKSSKDGVEIKHEGTTKVPKPKRVVLEYSF 68
 DB 26 KPFSLILGLIIVALLVAACGNTDSSKSSSTKDTISVKDENGTVKYPKDAKRIIVLEYSF 85
 QY 69 VDVALVDVKPVGIIADNKKRIIIPKLRDKIGKTVTSVGTQKPNLEISIKLPDLIADN 128
 DB 86 ADALVALDVKPVGIIADNKKRIIIPKVRKIGDVTSGTRKQPNLEISIKLPDLIADN 145
 QY 129 NRHKGIVKDLNKIAPTIELKSPDGYNENIDAFKTIISKALGKEEGKKLEEDHKIEY 188
 DB 146 SRHKGINKELNKIAPTIELKSPDGYNENIDAFKTIISKALNKEKEGKRLAEHDKLINKY 205
 QY 189 KKEITMDKNQKVLPAVAKSGLLAHPNSYVVGQFLSOLGPKFKAALSDVTKGLSKYKGPY 248
 DB 206 KDEIKFDRNQKVLPAVAKAGLLAHPNYSYVVGQFLNGLFKNALSDVTKGLSKYKGPY 265
 QY 249 LQMTTETLSQVNPFRMTNWKASNEPSLKELEKDPVWKKNVAVKQNRVILDRDLWR 308
 DB 266 LQDTEHLADLNPFRMTNWKASNEPSLKELEKDPVWKKNVAVKQNRVILDRDLWR 325
 QY 309 SRGLISSEEMAKELVELSKKQSK 331
 DB 326 SRGLISSEEMAKELVELSKKQSK 348
 RESULT 8

```

AAU33963
ID AAU33963 standard; protein; 309 AA.
XX
AC AAU33963;
XX
XX
DT 14-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation protein #239.
XX
XX
KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
KW antibacterial; drug design.
XX
XX
OS Staphylococcus aureus.
XX
XX
PN WO200170955-A2.
XX
XX
PD 27-SEP-2001.
XX
XX
PP 21-MAR-2001; 2001WO-US009180.
XX
PR 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
DR N-PSDB; AAS51822.
XX
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
XX
PS Example 3; SEQ ID NO 5459; 51lpp; English.
XX
XX
The invention relates to antisense inhibitors of genes essential to
prokaryotic cellular proliferation, their use in identifying the genes,
their use in the discovery of novel antibiotics, the essential genes
themselves and the encoded proteins. The prokaryotes used are Escherichia
coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
useful for the identification of potential new targets for antibiotic
development. The antisense nucleic acids can also be used to identify
proteins used in proliferation, to express these proteins, and to obtain
antibodies capable of binding to the expressed proteins. The proteins can
be used to screen compounds in rational drug discovery programmes. The
antisense nucleic acid sequence is also useful to screen for homologous
nucleic acids which are required for cell proliferation in a wide variety
of organisms. The present sequence represents an essential prokaryotic
cellular proliferation protein. Note: The sequence data for this patent
did not form part of the printed specification, but was obtained in
electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 309 AA;
XX
Query Match 69.9%; Score 1181; DB 4; Length 309;
Best Local Similarity 72.8%; Pred. No. 5.2e-88;
Matches 225; Conservative 38; Mismatches 46; Indels 0; Gaps 0;
XX
QY 16 LLFVLIAACGNNSSNSSKSSKDGVEIKHEGCTTKVPHKRVVVLVYSFVDAVAL 75
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1 LILVFLVAACGNTDNSSKSSSTKDTISVKDENGTVKVPDKAKRIVLVYSFADALAA 60
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 76 DVKPGVIADNKKRIIPLRDKIKGYTSVGRKOPNLEEISKLKPDLIADNNRHKGIV 135
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 61 DVKPGVIADGKKRIIPVREKIGNYTSVGRKOPNLEEISKLKPDLIADSSRHKGIN 120
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

```



```

SQ Sequence 319 AA;
Query Match 42.9%; Score 724.5; DB 6; Length 319;
Best Local Similarity 46.9%; Pred. No. 1.1e-50;
Matches 149; Conservative 63; Mismatches 95; Indels 11; Gaps 4;

QY 13 VIGLLFVLIATAACGNSSSSSSKSSKDGVEIKHEEGTTKVPKPKRVVVLVLEYSFVDAL 72
DB 9 VVFMILLVAVAGCCQKDT-----BEKTEMTTIKDELGTEKIKKPKRVVVLVLEYSFADYL 63

QY 73 VALDVKPGVIADNDKNRIIKPLDKIGKYTSVGRKQPNLEIEISKLPDLIIADNNRHK 132
DB 64 AALDNKPGVIADNDGSGTNIITKSVRDKIGAYESVGRPOPNMEVISKLPDLIIADVSRHK 123

QY 133 GIYKDLNKIAPTIELKSPDGYNNENIDAFKTIKALGKEEGKRLKLEHDKKIEYKKEI 192
DB 124 KIKSELSKIAPTIMLVSGTGYNNANIEAFKTVAKAVGKEKEKRLKLEHDKKILAEIRKKI 183

QY 193 TMDKNQKVLPAVAAGSGLLAHPSNSYVQFLSQLGFKALSDDVTGKLSKYLKGPYLOMN 252
DB 184 EQSLTKSAFAFGISRAGMFINNEDTFMGQFLIKMGIQPEVTKDTHVGERKGGFYIYN 243

QY 253 TETLSQVNPFRMFWTN-KASSNEPSLKELEKDP-VWKKLNAVKNQVRVDILDRDLWARSR 310
DB 244 NEELANINPKMILATDGTDKNRTKF----IDPAVWKSLLKAVKDKVYDVRNKLKSR 299

QY 311 GLISSEEMAKELVELSKK 328
DB 300 GIASESMAEDLEKIAEK 317

RESULT 13
ABR38895
ID ABR38895 standard; protein; 325 AA.
XX ABR38895;
AC ABR38895;
DT 10-MAY-2003 (first entry)
XX 10-MAY-2003 (first entry)
DE Surface exposed immunogenic polypeptide (SEIP) # SEQ ID 21.
XX SEIP; surface exposed immunogenic polypeptide; iron-binding; receptor;
KW antibacterial; fungicide; protozoacide; infection; disinfection.
XX Unidentified.
OS
XX WO2003006672-A2.
PN
XX 23-JAN-2003.
PD
XX 10-JUL-2002; 2002WO-US021663.
XX 10-JUL-2001; 2001US-0304390P.
PA (DSQU-) D-SQUARED BIOTECHNOLOGIES INC.
XX
PI Scott DL;
XX WPI; 2003-221756/21.
XX Isolating microbial target surface exposed immunogenic polypeptides, by
PT immunizing host with microbial membrane-associated polypeptides to
PT produce antibodies and using antibodies to isolate target polypeptides.
XX
XX Claim 15; Page 41; 46pp; English.
XX The invention relates to the isolation of target surface exposed
CC immunogenic polypeptides (SEIP) from microorganisms. The SEIP's of the
CC invention are receptors of iron-binding molecules. The activity of
CC polypeptides of the invention may be described as antibacterial,
CC fungicide and protozoacide. Pharmaceutical compositions based on the
CC polypeptides of the invention are useful for treating infections and
CC diseases caused by microbes including bacteria, fungi and protozoa in a
CC
CC patient. They are also useful for topical disinfection in an animal or
CC mammal e.g. human. The current sequence represents an SEIP of the
CC invention
XX
XX Sequence 325 AA;
Query Match 42.9%; Score 724.5; DB 6; Length 325;
Best Local Similarity 46.9%; Pred. No. 1.1e-50;
Matches 149; Conservative 63; Mismatches 95; Indels 11; Gaps 4;

QY 13 VIGLLFVLIATAACGNSSSSSSKSSKDGVEIKHEEGTTKVPKPKRVVVLVLEYSFVDAL 72
DB 15 VVFMILLVAVAGCCQKDT-----BEKTEMTTIKDELGTEKIKKPKRVVVLVLEYSFADYL 69

QY 73 VALDVKPGVIADNDKNRIIKPLDKIGKYTSVGRKQPNLEIEISKLPDLIIADNNRHK 132
DB 70 AALDNKPGVIADNDGSGTNIITKSVRDKIGAYESVGRPOPNMEVISKLPDLIIADVSRHK 129

QY 133 GIYKDLNKIAPTIELKSPDGYNNENIDAFKTIKALGKEEGKRLKLEHDKKIEYKKEI 192
DB 130 KIKSELSKIAPTIMLVSGTGYNNANIEAFKTVAKAVGKEKEKRLKLEHDKKILAEIRKKI 189

QY 193 TMDKNQKVLPAVAAGSGLLAHPSNSYVQFLSQLGFKALSDDVTGKLSKYLKGPYLOMN 252
DB 190 EQSLTKSAFAFGISRAGMFINNEDTFMGQFLIKMGIQPEVTKDTHVGERKGGFYIYN 249

QY 253 TETLSQVNPFRMFWTN-KASSNEPSLKELEKDP-VWKKLNAVKNQVRVDILDRDLWARSR 310
DB 250 NEELANINPKMILATDGTDKNRTKF----IDPAVWKSLLKAVKDKVYDVRNKLKSR 305

QY 311 GLISSEEMAKELVELSKK 328
DB 306 GIASESMAEDLEKIAEK 323

RESULT 14
ABR41867
ID ABR41867 standard; protein; 325 AA.
XX ABR41867;
AC ABR41867;
DT 11-AUG-2003 (first entry)
XX 11-AUG-2003 (first entry)
DE Staphylococcal surface-exposed immunogenic polypeptide D2 SA05.
XX Surface-exposed immunogenic polypeptide; SEIP; siderophore; receptor;
KW antibacterial; vaccine.
XX Staphylococcus aureus.
OS
XX WO2003020875-A2.
XX 13-MAR-2003.
XX 17-JUN-2002; 2002WO-US019224.
XX 17-JUN-2001; 2001US-0298975P.
PA (DSQU-) D-SQUARED BIOTECHNOLOGIES INC.
XX
PI Scott DL;
XX WPI; 2003-300870/29.
DB N-PSDB; ACC48534.
XX
XX Novel surface-exposed immunogenic polypeptide of Staphylococcus aureus
PT containing receptors for siderophores or iron-binding ligands, useful for
PT producing antibodies effective against Staphylococci infection.
XX
XX Claim 3; Page 55; 62pp; English.
XX The present sequence is that of a novel Staphylococcus aureus surface-
CC exposed immunogenic polypeptide (SEIP). To isolate SEIP genes, an

```


PA (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
PI N-PSDB; ACA47237.
DR WPI; 2003-029926/02.
XX N-PSDB; ACA47237.
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 71291; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 175 AA;
Query Match 38.1%; Score 644; DB 6; Length 175;
Best Local Similarity 70.9%; Pred. No. 1.8e-44;
Matches 124; Conservative 23; Mismatches 24; Indels 4; Gaps 1;
QY 5 VRGLKILSVIGLLFVLIAATACGNSSSNS----SKESKDGVEIKHEGTTKVPKPKR 60
DB 1 MKGFKPAGIVALLFALVLVLTACGNSGSGNSKSSKSIKIKHEIGTTKVPKDAKR 60
QY 61 VVLEYSFVDALVALDVKPGVADNKKRIIPLRDKTKGTSVCTRQPNLEISKLK 120
DB 61 VVLEFSFVDALAALNVKPGVADNKKPRIIPLRDKTKGTSVCTRQPNLEISKLK 120
QY 121 PLLIADNNRHGIYKDLNKAFTIELKSFDDGYNENIDAFKTIKALCKEKGK 175
DB 121 PLLIADNNRHGIYKELSKIAFTIELKSFDDGYNENIDAFKTIKALCKNDVGQ 175

RESULT 17
ABU38966
ID ABU38966 standard; protein; 298 AA.
XX
AC ABU38966;
XX
DT 19-JUN-2003 (first entry)
XX

DE Protein encoded by Prokaryotic essential gene #24493.
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX Pasteurella multocida.
XX WO200277183-A2.
XX 03-OCT-2002.
XX 21-MAR-2002; 2002WO-US009107.
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
DR WPI; 2003-029926/02.
XX N-PSDB; ACA42836.
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 66890; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 298 AA;
Query Match 30.6%; Score 516.5; DB 6; Length 298;
Best Local Similarity 38.1%; Pred. No. 9.5e-34;
Matches 111; Conservative 65; Mismatches 98; Indels 17; Gaps 4;
QY 43 VEIKHEGTTKVPKPKRVVLEYSFVDALVALDVKPGVADNKKRIIPLRDKIGKY 102
DB 23 VTVKQKGFTLDSVPKRVVLEYSVDALAIGVSPGVGADNDKTRILQKVRDKVQPW 82

the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid; (2) a host cell containing the vector; (3) an isolated encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 300 AA;

Query Match 27.2%; Score 460; DB 6; Length 300;
 Best Local Similarity 35.8%; Pred. No. 4e-29;
 Matches 101; Conservative 60; Mismatches 105; Indels 16; Gaps 4;
 QY 45 IKHEGTTKVPKPRVVLVLEYSFVDALVALDVPKVGADDDNKKRIIPLRDKIGKYS 104
 Db 24 VDEHGTFTLEKTPQRIIVVLEISFADALAAVDVPIGIIADDDNDAKRIILPEVRAHLKPMQS 83
 QY 105 VQTRKQPNLEISKLKPDLLIADNNRHGIYKLNKIAPTIELKSFQDGYNENIDAFKTI 164
 Db 84 VQTRAQPSLEATAALKPDLLIADSSRHAGIYVIALQOIAPVLLKSRNETHYAEHLASAII 143
 QY 165 SKALGKEEGKKRLEHDKKIEYKKEITMDKNQKVLPAVAAGSGLLAHPSNSYVQGFUS 224
 Db 144 GEVVGKKREMQARLEQHKEMQWSSQ--LPTGTRVAFGTSREQLFNLHTQETWTGSLA 201
 QY 225 QLGFKREALSDVTKGLSKYLKGPYL-QMNTETLSQVNPMPFIMTNKASSNPSPKLEBK 283
 Db 202 SLGL-----NVPAMAGASMPISGLEQLLAVNPAPLLVAHYR---EESIVKRWQ 248
 QY 284 DPVWKKLNKQVQVVDLDRDLWARSRLISSEEMAKELVEL 325
 Db 249 DPLWQMLTAAQKQVQVSDNSNWARMRGIFAAERIAADTVKI 290

RESULT 20

AAU34876
 ID AAU34876 standard; protein; 302 AA.

XX

AC AAU34876;

XX 14-FEB-2002 (first entry)

DT

XX E. coli cellular proliferation protein #457.

XX Antisense; prokaryotic cellular proliferation protein; antibiotic;

KW antibacterial; drug design.

XX

OS Escherichia coli.

XX

PN WO200170955-A2.

XX

PD 27-SEP-2001.

XX

PF 21-MAR-2001; 2001WO-US009180.

XX

PR 21-MAR-2000; 2000US-0191078P.

PR 23-MAY-2000; 2000US-0206848P.

PR 26-MAY-2000; 2000US-0207727P.

PR 23-OCT-2000; 2000US-0242578P.

PR 27-NOV-2000; 2000US-0253625P.

PR 22-DEC-2000; 2000US-0257931P.

PR 16-FEB-2001; 2001US-0269308P.

XX (ELIT-) ELITRA PHARM INC.

XX

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

PI Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

DR N-PSDB; AAS52735.

XX

XX New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids.

XX

XX Example 3; SEQ ID NO 10469; 511pp; English.

XX

XX The invention relates to antisense inhibitors of genes essential to

CC prokaryotic cellular proliferation, their use in identifying the genes,

CC their use in the discovery of novel antibiotics, the essential genes

CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*

CC *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,

CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also

CC useful for the identification of potential new targets for antibiotic

CC development. The antisense nucleic acids can also be used to identify

CC proteins used in proliferation, to express these proteins, and to obtain

CC antibodies capable of binding to the expressed proteins. The proteins can

CC be used to screen compounds in rational drug discovery programmes. The

CC antisense nucleic acid sequence is also useful to screen for homologous

CC nucleic acids which are required for cell proliferation in a wide variety

CC of organisms. The present sequence represents an essential prokaryotic

CC cellular proliferation protein. Note: The sequence data for this patent

CC did not form part of the printed specification, but was obtained in

CC electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 302 AA;

Query Match 27.2%; Score 460; DB 4; Length 302;

Best Local Similarity 35.5%; Pred. No. 4e-29;

Matches 100; Conservative 62; Mismatches 104; Indels 16; Gaps 4;

QY 45 IKHEGTTKVPKPRVVLVLEYSFVDALVALDVPKVGADDDNKKRIIPLRDKIGKYS 104

Db 26 VDEHGTFTLEKTPQRIIVVLEISFADALAAVDVPIGIIADDDNDAKRIILPEVRAHLKPMQS 85

QY 105 VQTRKQPNLEISKLKPDLLIADNNRHGIYKLNKIAPTIELKSFQDGYNENIDAFKTI 164

Db 86 VQTRAQPSLEATAALKPDLLIADSSRHAGIYVIALQOIAPVLLKSRNETHYAEHLASAII 145

QY 165 SKALGKEEGKKRLEHDKKIEYKKEITMDKNQKVLPAVAAGSGLLAHPSNSYVQGFUS 224

Db 146 GEVVGKKREMQARLEQHKEMQWSSQ--LPTGTRVAFGTSREQLFNLHTQETWTGSLA 203

QY 225 QLGFKREALSDVTKGLSKYLKGPYL-QMNTETLSQVNPMPFIMTNKASSNPSPKLEBK 283

Db 204 SLGL-----NVPAMAGASMPISGLEQLLAVNPAPLLVAHYR---EESIVKRWQ 250

QY 284 DPVWKKLNKQVQVVDLDRDLWARSRLISSEEMAKELVEL 325

Db 251 DPLWQMLTAAQKQVQVSDNSNWARMRGIFAAERIAADTVKI 292

```
RESULT 21
ABU28881
ID ABU28881 standard; protein; 302 AA.
XX AC ABU28881;
XX DT 19-JUN-2003 (first entry)
XX DE Protein encoded by Prokaryotic essential gene #14408.
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX OS Escherichia coli.
XX PN WO200277183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX FA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX PI Wall D, Tamwick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX DR WPI; 2003-029926/02.
XX DR N-PSDB; ACA32751.
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX PS Claim 25; SEQ ID NO 56805; 1766pp; English.
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX CC proliferation or the activity of a gene in an operon required for
XX CC proliferation; (7) identifying a compound that influences the activity of
XX CC the gene product or that has an activity against a biological pathway
XX CC required for proliferation, or that inhibits cellular proliferation; (8)
XX CC identifying a gene required for cellular proliferation or the biological
XX CC pathway in which a proliferation-required gene or its gene product lies
XX CC or a gene on which the test compound that inhibits proliferation of an
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX CC compound's activity; (11) a culture comprising strains in which the gene
XX CC product is overexpressed or underexpressed; (12) determining the extent
XX CC to which each of the strains is present in a culture or collection of
XX CC strains; or (13) identifying the target of a compound that inhibits the
XX CC proliferation of an organism. The antisense nucleic acids are useful for
XX CC identifying proteins or screening for homologous nucleic acids required
XX CC for cellular proliferation to isolate candidate molecules for rational
XX CC drug discovery programs, or for screening homologous nucleic acids
XX CC required for proliferation in cells other than S. aureus, S. typhimurium,
XX CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
XX CC the target prokaryotic essential genes. Note: the sequence data for this
XX CC patent did not form part of the printed specification, but was obtained
XX CC in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX CC Sequence 302 AA;
XX
```

```
Query Match 27.2%; Score 460; DB 6; Length 302;
Best Local Similarity 35.5%; Pred. No. 4e-29;
Matches 100; Conservative 62; Mismatches 104; Indels 16; Gaps 4;
QY 45 IKHEEGTTKVPKPKRVVILEYSFVDALVALDVKVPVGADNNKKNRIKPLRDKIGKYTS 104
DB 26 VQDEHGTFLEKTPQRIIVLELSPADALAAVDVIPIGIADDNDKAKILPEVRAHLKPQWS 85
QY 105 VGRKQKNLBEISKLPDLIIADNNRHKGIIKDLINKIAPTIELKSFQDGYNNENIDAFKTI 164
DB 86 VGTQAQPSLEAIAALKPDLIADSSRHAGVYIALQQIAPVLLKSRNETYAENLQSAAIL 145
QY 165 SKALGKEEGKRLLEHDKKIEEYKKEITMDKNOKVLPAAVAKSGLLAHPNSVYQFSL 224
DB 146 GEMVKKREMQARLEQHKERMAQWASQ--LPKGRVAFGTSREQQFNLIHTQETWTGSLA 203
QY 225 QLGFEALSDDVTKGLSKYLKGPYL-QMNTETLSQVNPFRPIMTNKASSNEPSLKELEK 283
DB 204 SLGL-----NVPAAMAGASMPISIGLEQLLAVNPALLVAHYR---EESIVKRWQQ 250
QY 284 DPVWKLNNAVKNQVRIIDRLDRLWARSGLISSSEMAKELVEL 325
DB 251 DPLWQMLTAAQKQOVASVDSNTWARMRGIFAERIAADTVKI 292
RESULT 22
ABU41496
ID ABU41496 standard; protein; 297 AA.
XX AC ABU41496;
XX DT 19-JUN-2003 (first entry)
XX DE Protein encoded by Prokaryotic essential gene #27023.
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX OS Pseudomonas syringae.
XX PN WO200277183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX FA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX PI Wall D, Tamwick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX DR WPI; 2003-029926/02.
XX DR N-PSDB; ACA45366.
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX PS Claim 25; SEQ ID NO 69420; 1766pp; English.
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
XX CC antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
```

antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 297 AA;

Query Match 26.9%; Score 454; DB 6; Length 297;

Best Local Similarity 34.3%; Pred. No. 1.2e-28;

Matches 109; Conservative 60; Mismatches 109; Indels 40; Gaps 6;

QY 17 LFLVLTATACGNNSSNSSKESKDGVEIKHEEGTTKV--PKHPRVVVLEYSFVDALVA 74
 DB 5 LTLTASAA-----QAAPIDDDGQHKVHLPTPKRVVLEFSFDGLAS 49
 QY 75 LDVQVPGIADNKNRIIKPLRDKIGKVTSGTRKQPNLEETSKLPDLIIADNNRHKG 134
 DB 50 VGVTPGAADDGASRVLPKRVKAVGWSGLRGQPNIEVTLARKLPDLIIADLGRHQA 109
 QY 135 YKDLNKIAPTIELKSGFDGYNENIDAFKTI SKALGKEEGKRLBEHDKKIEYKEITM 194
 DB 110 YNDLASLAPTLPSRGEDYQGLSKSAGLIGMALGKGPQMAEIAENRQHLTVAEQIPA 169
 QY 195 DGNQKVLPAVAKSGLLAHPNSYVQFLSQFLSGFKEALSDDVTYKGLSKYKLPYLMQTE 254
 DB 170 DSN--VLFVGVADEDSFVGHPSYAGSVLQAIG-----LQVPEVRNNA 211
 QY 255 TLSQVNPFRMFTWTK-----ASSNEPSLKEL-EKDPVWKKNLNAVKNQVVDILDRDLWARS 309
 DB 212 PTEFVSLEQLLADPNWLLVGHYRPSIVDTWSKQLWQLVGNVKNQVAEVDGDSWARN 271
 QY 310 RGLISEEMAKELVELSK 327
 DB 272 RGMASEQIADALAVLK 289

RESULT 23

AAW89745

ID AAW89745 standard; protein; 330 AA.

AAW89745;

16-MAR-1999 (first entry)

Staphylococcus aureus protein SEQ ID #5193.

Computer readable medium; vaccine; *S. aureus* infection; immunodetection; cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy; skin infection; surgical wound infection; scalded skin syndrome; toxic shock syndrome.

Staphylococcus aureus.

PN EP786519-A2.

XX 30-JUL-1997.

XX 07-JAN-1997; 97EP-00100117.

XX 05-JAN-1996; 96US-0009861P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA;

XX WPI; 1997-374922/35.

XX Polynucleotide(s) and proteins derived from *Staphylococcus aureus* -

XX stored on computer readable medium and used in the production of anti-

XX *S. aureus* vaccines.

XX Claim 23; Page 3189-3190; 3271pp; English.

XX This sequence represents a *Staphylococcus aureus* protein sequence of the

XX invention. The DNA sequences encoding the *S. aureus* proteins are recorded

XX on a computer readable medium, preferably selected from a floppy or hard

XX disk, random access memory (RAM), read-only memory (ROM) or CD-ROM.

XX Homology searches using the *S. aureus* DNA sequences allows putative

XX functions to be assigned so that protein-encoding or regulatory regions

XX of commercial, therapeutic or industrial importance can be obtained.

XX Specifically, sequences which are likely to encode antigens have been

XX identified and these polypeptides can be used in a vaccine composition

XX against *S. aureus* infection. The polypeptides can also be used in a kit

XX for the immunodetection of *S. aureus* in a sample. *S. aureus* is implicated

XX in numerous human diseases, including cellulitis, eyelid infections, food

XX poisoning, osteomyelitis, skin and surgical wound infections, scalded

XX skin syndrome, toxic shock syndrome, etc. Organisms transformed with the

XX DNA sequences can be used for recombinant production of the polypeptides.

XX The new DNA sequences (and their fragments) are useful as primers or

XX probes for isolating homologues of any of the 5191 *S. aureus* DNA sequences

XX contained on the computer readable medium

XX Sequence 330 AA;

QY Query Match 24.0%; Score 405; DB 2; Length 330;

DB Best Local Similarity 32.8%; Pred. No. 1.4e-24;

QY Matches 113; Conservative 65; Mismatches 132; Indels 34; Gaps 12;

QY 8 LKILSVIGLLFVLITATAACGNNSSNSSKESKDGVEIKHEEGTTKVPKHKRVVTVLEYS 67

DB 5 IKML-VVTLAFLV-LAGCGSNKQSSDNKDKTTSIKHAMGTTTEIKGPKRVVTVLYG 62

QY 68 FVDALVALDVKPVGIADNKNRIIKPLRDKIGKVTSGTRKQPNLEETSKLPDLIIAD 127

DB 63 ATDVAVSLGVKPVGAVESWTQPKFEYIKNDLKDTKIVGOEPAPNLEETSKLPDLIVAS 122

QY 128 NNRHKGIVKDLNKIAPTIELKSGFDGYNENIDAFKTI SKALGKEEGKRLBEHDKKIEE 187

DB 123 KVRNEKVVYDQLSKIAPTIV---STDVTFKFK-DTTKMGKALGKEAEADLLKKYDDKVA 178

QY 188 YKKEITMDKNQKVLPAVAKSGLLAHPN-----SYVQFLSQFLGKEALSDDVTYKGL 240

DB 179 FQK----DAKAKYKDAWPLKASVWNFRADHTRIYAGGAGETLNDLGFKR-----NKDL 228

QY 241 SKYLAG--PYLQMTN-ETLSOVNPERMFTWTKASSNEPSL-----KELEKDPVWKKNLA 292

DB 229 QKQVNDGKDIQTLSKESIPLMNADHIFVVKSDPNKADALVKKYTESWTSSEWKNLDA 288

QY 293 VKNQV--DILDRDLWARSGLISSEEMAKELVELSKDKSKDNK 335

DB 289 VKNQVSDDLDEITWNLAGGYSKLLIDDLVE--KLNIEKQSK 330

RESULT 24

AAW42331

ID AAW42331 standard; protein; 330 AA.

XX SEIP; surface exposed immunogenic polypeptide; iron-binding; receptor;
 KW antibacterial; fungicide; protozoacide; infection; disinfection.
 XX Unidentified.
 OS WO2003006672-A2.
 XX 23-JAN-2003.
 XX 10-JUL-2002; 2002WO-US021663.
 XX 10-JUL-2001; 2001US-0304390P.
 XX (DSQU-) D-SQUARED BIOTECHNOLOGIES INC.
 XX Scott DL;
 XX WPI; 2003-221756/21.
 XX Isolating microbial target surface exposed immunogenic polypeptides, by
 PT immunizing host with microbial membrane-associated polypeptides to
 PT produce antibodies and using antibodies to isolate target polypeptides.
 XX Claim 15; Page 40; 46pp; English.
 XX The invention relates to the isolation of target surface exposed
 CC immunogenic polypeptides (SEIP) from microorganisms. The SEIP's of the
 CC invention are receptors of iron-binding molecules. The activity of
 CC polypeptides of the invention may be described as antibacterial,
 CC fungicide and protozoacide. Pharmaceutical compositions based on the
 CC polypeptides of the invention are useful for treating infections and
 CC diseases caused by microbes including bacteria, fungi and protozoa in a
 CC patient. They are also useful for topical disinfection in an animal or
 CC mammal e.g. human. The current sequence represents an SEIP of the
 CC invention
 XX Sequence 330 AA;
 SQ
 Query Match 24.0%; Score 405; DB 6; Length 330;
 Best Local Similarity 32.8%; Pred. No. 1.4e-24;
 Matches 113; Conservative 65; Mismatches 132; Indels 34; Gaps 12;
 QY 8 LKILSVIGLLPVLIIATAACGNNSSSSKSSKDGVEIKHEEGTTKVPKPKRVVVLVLEYS 67
 DB 5 IKML-VVTLAFLV-LAGCGSGNSKQSSDNKDKETTSIKHAMGTTTEIKGPKRVVTLVYQ 62
 QY 68 FVDALVALDVKPGVIADDNKNRIKPLRDKIGKYSVGRKQPNLEISKLKPDLLIAD 127
 DB 63 ATDVAVSLGVFPVGAVESWTQPKPEYIKNDLKTKIVGQBPAPNLEISKLKPDLLIVAS 122
 QY 128 NNRHKGIIYKDLNKTAPTIELKSPFDGYNENIDAFKTSIKALGKEEGKRLLEHDKKIEE 187
 DB 123 KVRNEKYVDQLSKIAPTIV---STDVTFKFK-DTTKLMGKALGKEAEDLLKKYDDKVA 178
 QY 188 YKKEITMDKNQKVLPAVAKSGLLAHPN-----SYVGQFLSQLGKEALSDDVTGKL 240
 DB 179 FQK----DAKAKYKDAWPLKASVNFRAHDTRIYAGGYAGIILNDLGFKR-----NKDL 228
 QY 241 SKYLGK--PYLQMT--ETLSQVNPFRFIMTNKASSNEPSL-----KELEKDPVWKKLNA 292
 DB 229 QKQVNDGKDIQLTSKESIPLMNADHIFVVKSDPNKADALVKKTESWTSKKEWKNLDA 288
 QY 293 VKNQRV--DILDRDLWARGSLISSEEMAKELVELSKDKSKDNK 335
 DB 289 VKNQVSDDLDEITWNLAGGYSKSLKLDLDE--KLNIEKQSK 330
 RESULT 28
 ABR41864
 ID ABR41864 standard; protein; 330 AA.
 XX
 AC ABR41864;

XX 11-AUG-2003 (first entry)
 XX Staphylococcal surface-exposed immunogenic polypeptide D2 SA02.
 DE Surface-exposed immunogenic polypeptide; SEIP; siderophore; receptor;
 KW antibacterial; vaccine.
 KW Staphylococcus aureus.
 OS WO2003020875-A2.
 XX 13-MAR-2003.
 XX 17-JUN-2002; 2002WO-US019224.
 XX 17-JUN-2001; 2001US-0298975P.
 XX (DSQU-) D-SQUARED BIOTECHNOLOGIES INC.
 XX Scott DL;
 XX WPI; 2003-300870/29.
 XX N-PSDB; ACC48531.
 XX Novel surface-exposed immunogenic polypeptide of Staphylococcus aureus
 PT containing receptors for siderophores or iron-binding ligands, useful for
 PT producing antibodies effective against Staphylococci infection.
 XX Claim 3; Page 51; 62pp; English.
 XX The present sequence is that of a novel Staphylococcus aureus surface-
 CC exposed immunogenic polypeptide (SEIP). To isolate SEIP genes, an
 CC expression library of S. aureus genomic DNA was screened with anti-SEIP
 CC antisera. DNA was isolated from clones expressing SEIPs, and amino acid
 CC sequences were deduced and analysed for conserved and/or functional
 CC domains. The present SEIP has sequence homology with the siderophore
 CC family of periplasmic binding proteins. The invention provides methods
 CC for the identification, production and recovery of SEIPs. The SEIPs can
 CC be used individually, or in combination, to produce anti-staphylococcal
 CC antibodies useful in passive or active immunisation strategies to prevent
 CC or contain staphylococcal infection. They can also be used to develop
 CC diagnostic assays
 XX Sequence 330 AA;
 SQ
 Query Match 24.0%; Score 405; DB 6; Length 330;
 Best Local Similarity 32.8%; Pred. No. 1.4e-24;
 Matches 113; Conservative 65; Mismatches 132; Indels 34; Gaps 12;
 QY 8 LKILSVIGLLPVLIIATAACGNNSSSSKSSKDGVEIKHEEGTTKVPKPKRVVVLVLEYS 67
 DB 5 IKML-VVTLAFLV-LAGCGSGNSKQSSDNKDKETTSIKHAMGTTTEIKGPKRVVTLVYQ 62
 QY 68 FVDALVALDVKPGVIADDNKNRIKPLRDKIGKYSVGRKQPNLEISKLKPDLLIAD 127
 DB 63 ATDVAVSLGVFPVGAVESWTQPKPEYIKNDLKTKIVGQBPAPNLEISKLKPDLLIVAS 122
 QY 128 NNRHKGIIYKDLNKTAPTIELKSPFDGYNENIDAFKTSIKALGKEEGKRLLEHDKKIEE 187
 DB 123 KVRNEKYVDQLSKIAPTIV---STDVTFKFK-DTTKLMGKALGKEAEDLLKKYDDKVA 178
 QY 188 YKKEITMDKNQKVLPAVAKSGLLAHPN-----SYVGQFLSQLGKEALSDDVTGKL 240
 DB 179 FQK----DAKAKYKDAWPLKASVNFRAHDTRIYAGGYAGIILNDLGFKR-----NKDL 228
 QY 241 SKYLGK--PYLQMT--ETLSQVNPFRFIMTNKASSNEPSL-----KELEKDPVWKKLNA 292
 DB 229 QKQVNDGKDIQLTSKESIPLMNADHIFVVKSDPNKADALVKKTESWTSKKEWKNLDA 288
 QY 293 VKNQRV--DILDRDLWARGSLISSEEMAKELVELSKDKSKDNK 335
 DB 289 VKNQVSDDLDEITWNLAGGYSKSLKLDLDE--KLNIEKQSK 330

RESULT 29
ADF43556
ID ADF43556 standard; protein; 330 AA.
XX
AC ADF43556;
XX
DT 12-FEB-2004 (first entry)
XX
DE Staphylococcus aureus cbrA.
XX
KW antibacterial; immunostimulant; vaccine; vaccine;
KW Staphylococcus aureus infection; infection prevention;
KW infection attenuation; cbrA.
XX
OS Staphylococcus aureus.
XX
PN US2003153733-A1.
XX
PD 14-AUG-2003.
XX
PF 24-OCT-2002; 2002US-00278946.
XX
PR 20-MAR-1998; 98US-0078862P.
PR 01-APR-1998; 98US-0080296P.
PR 07-MAY-1998; 98US-0084674P.
PR 18-MAR-1999; 99WO-US006199.
PR 15-JAN-2002; 2002US-00830217.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Simpson AJG, Choi GH;
XX
DR WPI; 2003-897737/82.
DR N-PSDB; ADF43555.
XX
PT New Staphylococcus aureus gene, useful for preparing a vaccine against
PT infection caused by Staphylococcus aureus.
XX
PS Claim 9; SEQ ID NO 16; 61pp; English.
XX
CC The invention describes a new isolated nucleic acid molecule comprising:
CC a sequence that is 95% identical to a fully defined sequence having 586-
CC 2226 bp, or its complement; or a sequence encoding a polypeptide having a
CC fully defined sequence comprising 136-691 amino acids, or its complement.
CC The gene is Staphylococcus aureus gene. The nucleic acid is useful for
CC preparing a vaccine against infection caused by Staphylococcus aureus.
CC The methods are useful for preventing or attenuating an infection caused
CC by a Staphylococcus, detecting Staphylococcus nucleic acids in a
CC biological sample obtained from an animal, and detecting Staphylococcus
CC antibodies in a biological sample obtained from an animal. This is the
CC amino acid sequence of Staphylococcus aureus cbrA protein.
XX
SQ Sequence 330 AA;
Query Match 24.0%; Score 405; DB 7; Length 330;
Best Local Similarity 32.8%; Pred. No. 1.4e-24;
Matches 113; Conservative 65; Mismatches 132; Indels 34; Gaps 12;
QY 8 LKILSVIGLFLVLIATAACGNSSNSSKESKGVIEIKHEGTTKVPKPKRVVVLEYS 67
Db 5 IKGL-VTLAFILV-LAGCGSNKQSSNDKKTTSIKHMGTTIEIKGPKRVVTLVQG 62
QY 68 FVDALVALDVKPVGIADNKNRIIIPKLRDKIGKTVSVGRKQPNLEEISKLPLIITAD 127
Db 63 ATDVAVSLGVKPVGAVESWTQPKFVIKNDLKDTKIVGEPAPNLEEISKLPLIIVAS 122
QY 128 NNRHGIYKDLNKAIPITIELKSPDGDYNNIDAFKTIKALKESEKGRLEEHKKEE 187
Db 123 KVRNEKVDQLSKIAPTV---STDVTVPFK-DTTKLMGKALKEAEADLLKKYDDKVA 178
QY 188 YKKEITMDKNQKVLPAVAAKSGLLAHPSN-----SYVGQFLSQLGFKREALSDVDVTKGL 240

Db 179 FOK-----DAKAKYKDAWPLKASVVNFRADHTRIYAGGYAGEILNDLGFKR-----NKDL 228
QY 241 SKYLGK--PYLQMT--ETLSQVNPFRMFIIMTKASSNEPSL-----KELEKDPVWKKLNA 292
Db 229 QKQVDNGKDIITLTSKSGSIPLMNADHIFVVKSDPNADKDALVVKTESEWTSSKEWKNLDA 288
QY 293 VKNQRV--DILDRDLWARSRLISSEEMAKELVELSKDKSKDNK 335
Db 289 VKNQVSDDLDEITWNLGGYKSSKLIDDIYE--KLNIEKQSK 330
RESULT 30
ABU17675
ID ABU17675 standard; protein; 324 AA.
XX
AC ABU17675;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by prokaryotic essential gene #3202.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Bacillus anthracis.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA21545.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 45599; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway;
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the

CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 324 AA;

Query Match 23.1%; Score 391; DB 6; Length 324;
 Best Local Similarity 34.4%; Pred. No. 1.9e-23;
 Matches 116; Conservative 59; Mismatches 122; Indels 40; Gaps 13;
 QY 9 KILSVIGLLFVLITAAAGNNSSNSSKESKD---GVEIKHEEGTKVPKPKRVVVL 64
 DB 3 KILSI---FIVVFLFVAGCGQKEKETKADNNQQAITIKHAGETKLDKPAKKVVVL 58
 QY 65 EYSFVDALVALDVKPVGADDNKNRII---KPLRDKIGKYTSVGTQKQNLBEISKLK 120
 DB 59 EWWYSEDLALGVQPVGMADIKYNNKWNTKTSKQVV---DVGTRQPNLEISRLK 114
 QY 121 PDLIADNNRHGIYKDLNKIAPTIEL---KSPGDYNNENIDAKTISKALGKKEEGKKR 177
 DB 115 PDLIITTSFRGKAINKLEIQIAPTVMFDPSTSNNDHFAEMTETFKQIAKAVGKEEGKKV 174
 QY 178 LEEHDKKIEEYKKEI---TMDKNQKVLPAVAKS---GLLAHPSNSYVQGFLSQLGFK 229
 DB 175 LADMKAFAADAKAKIEKADLKDKNIAQAFTAKNVPTFRILT--DNSLALQVTKKLG-- 230
 QY 230 EALSDDVTGKLSKYLKGPYLQWNTETLSQVNPFR-MFMTNKAASSNEPSLKELEKDPVWK 288
 DB 231 --LNTPEAGKSE--PDGFKTTVESLQSVQDSNFYIIVADEDNIDFTQLK--GNPAWE 283
 QY 289 KLNVAQNQRVDILDRDLWARSRLISSEEMAKELVEL 325
 DB 284 ELKPKKENMYKLKGDWTFIFG-GPESATSLATQVADV 319

RESULT 31

ABG08851
 ID ABG08851 standard; protein; 1132 AA.

XX AC ABG08851;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #8942.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PP 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX XX (HYSE-) HYSEQ INC.

XX XX Drmanac RT, Liu C, Tang YT;

XX XX WPI, 2001-639362/73.

XX DR N-P8DB; AAS73038.

XX

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX Claim 20; SEQ ID NO 39210; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 1132 AA;

Query Match 22.2%; Score 375; DB 4; Length 1132;
 Best Local Similarity 32.5%; Pred. No. 2e-21;
 Matches 96; Conservative 52; Mismatches 115; Indels 32; Gaps 5;
 QY 45 IKHEEGTKVPKPKRVVVLSEYSPVDALVALDVKPVGADDNKNRIIKPLRDKIGKYTS 104
 DB 24 VQDEHGTILTLEKTPQRIIVLELSFADALAAVDVIPGIADNDKAKRILPEVRAHLKPQWS 83
 QY 105 VGTQKQNLBEISKLKPDIIADNNRHGIYKDLNKIAPTIELSKSPGDYNNENIDAKFTI 164
 DB 84 VGTQKQNLBEISKLKPDIIADNNRHGIYKDLNKIAPTIELSKSPGDYNNENIDAKFTI 143
 QY 165 SKALGKEEGKKRLKEHDKKIEEYKKEITMDKNQKVLPAVAKSGLLAHPSNSYVQGFLS 224
 DB 144 GEMVKKREMQARLEQHKERMAQWASQ--LPKGTAVAFGTSREQQFNLHTQETWTGSLA 201
 QY 225 QLGFKEALSDVDTGKLSKYLKGPYLQWNTETLSQVNPFRMFMNTKASSNEPSLKELEKD 284
 DB 202 SLGL-----NVPAAAGASNPSPGPGAGCPRCPSIHAITVTSITPCL----- 245
 QY 285 PVWKKLNAVKNQRVDIL-----DRDLWARSRLISSE-----EMAKELVELSK 327
 DB 246 --WTPGSVIRCRPVGVLKMTDEAGEDAKLVAVPHSKLSKEYDHIKDVNDLPPELLK 298

RESULT 32

ABG13148
 ID ABG13148 standard; protein; 932 AA.

XX AC ABG13148;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #13139.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX

```
PD 11-OCT-2001.
XX
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Dmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
DR N-PSDB; AAS77335.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 20; SEQ ID NO 43507; 103pp; English.
PS
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 932 AA;
SQ
Query Match 20.7%; Score 350; DB 4; Length 932;
Best Local Similarity 46.6%; Pred. No. 1.7e-19;
Matches 69; Conservative 33; Mismatches 46; Indels 0; Gaps 0;
QY 45 IKHEGGTTKVPKHKRVVLEYSFVDALVALDVKPVGIADNKKRIIKPLRDKIGKYTS 104
DB 558 VODEHGTFTLEKTPORIVVLELSFADALAADVPIGIGIADNNDKRIPEVRAHKLPQOS 617
QY 105 VGTBQPNLEEISKLKPDLLIADNNRHKGIIYKDLNKAFTIELKSPDGDYNNIDAFKTI 164
DB 618 VGTBQPSLEATAALKPDLLIADSSRHAGVYIALQQAIPVLLKSRNETYAENLQSAAT 677
QY 165 SKALGKEEGKRLKLEHDKKIEYKKEI 192
DB 678 GEMVGKKREMQARLEQHKERMAQWASQL 705
RESULT 33
ABM73050
ID ABM73050 standard; protein; 287 AA.
XX
XX ABM73050;
AC
XX
XX 20-NOV-2003 (first entry)
DT
XX
XX Staphylococcus aureus protein #2290.
DE
XX
XX Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
KW
```

```
enzymatic assay; antibiotic target.
KW
XX Staphylococcus aureus.
OS
XX
XX WO200294868-A2.
FN
XX
XX 28-NOV-2002.
PD
XX
XX 27-MAR-2002; 2002WO-IB002637.
PF
XX
XX 27-MAR-2001; 2001GB-00007661.
XX
XX (CHIR-) CHIRON SPA.
PA
XX
XX Masignani V, Mora M, Scarselli M;
PI
XX WPI; 2003-120786/11.
XX
XX N-PSDB; ACF74610.
DR
XX
XX New Staphylococcus aureus protein, useful as a vaccine for treating or
PT preventing Staphylococcal infection, specifically an infection caused by
PT S. aureus, e.g. sepsis.
XX
XX Claim 1; SEQ ID NO 4580; 49pp; English.
PS
XX The invention relates to novel genes and encoded proteins from
CC Staphylococcus aureus. A composition comprising the S. aureus protein, a
CC nucleic acid encoding the protein, or an antibody to the protein, is
CC useful as a pharmaceutical, particularly as a vaccine for treating or
CC preventing infection due to Staphylococcus bacteria, specifically an
CC infection caused by S. aureus. The composition is particularly useful for
CC treating or preventing sepsis in a patient. The composition can also be
CC used for diagnostics. The protein is also used in an assay for enzymatic
CC studies and as a target for antibiotics. This sequence represents one of
CC the novel S. aureus proteins of the invention
XX
XX Sequence 287 AA;
SQ
Query Match 20.3%; Score 343; DB 6; Length 287;
Best Local Similarity 32.1%; Pred. No. 1.4e-19;
Matches 97; Conservative 58; Mismatches 115; Indels 32; Gaps 10;
QY 50 GTTVPKHKRVVLEYSFVDALVALDVKPVGIADNKKRIIKPLRDKIGKYTSVGTGRK 109
DB 2 GTTEIKGKPKRVVTLVYQCATDVAVSLGVKPVGAVESWTQKPKFEYKNDLKDXTKIVGQEP 61
QY 110 QPNLEEISKLKPDLLIADNNRHKGIIYKDLNKAFTIELKSPDGDYNNIDAFKTIKALG 169
DB 62 APNLEEISKLKPDLLIVASKVRNEKYVDLSKIAPT---STDTVFKFK-DTTKLMGKALG 117
QY 170 KEEGKKELEHDKKIEYKKEITMDKNQKVLPAVAAKSGLLAHPSN-----SYVGQF 222
DB 118 KEKAEDLLKKYDDKVAAFQK----DAKAKYDAPLKAASVNFADHTRIVAGGYAGEI 173
QY 223 LSQLGFEKALSDVVTGKLSKYLKG--PYLQMT-ETLSQVNPFRMFINTKASSNEPSL- 278
DB 174 LNDLGFKR-----NKDLQKQVDNGKDIQLTSKESIPLMNADHIFVVKSDPNKDAALV 227
QY 279 ----KELEKDPVKKLNKAVKQV--DILDRDLWARSGLISSEENAKELVELSKDKSKD 333
DB 228 KKTSEWTSKKEWKNLDAVKNQVSDDLDEITWNLAGGYKSLKLIDDLYE--KLNIEKQ 285
QY 334 NK 335
DB 286 SK 287
RESULT 34
AAG90121
ID AAG90121 standard; protein; 315 AA.
XX
XX AAG90121;
AC
XX
```

DT 26-SEP-2001 (first entry)
 DE C glutamicum protein fragment SEQ ID NO: 3875.
 XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis.
 XX Corynebacterium glutamicum.
 OS EP1108790-A2.
 PN 20-JUN-2001.
 PD 18-DEC-2000; 2000EP-00127688.
 PF 16-DEC-1999; 99JP-00377484.
 PR 07-APR-2000; 2000JP-00159162.
 PR 03-AUG-2000; 2000JP-00280988.
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX WPI; 2001-376931/40.
 DR N-PSDB; AAH65340.
 XX Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX Claim 17; SEQ ID NO 3875; 246pp + Sequence Listing; English.
 PS The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and analyzing
 CC the expression profile or expression pattern of a gene derived from
 CC Coryneform bacterium, and identifying a homologue of a gene derived from
 CC coryneform bacterium. Coryneform bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a protein described in the
 CC exemplification of the invention. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from the European Patent Office
 XX Sequence 315 AA;
 SQ
 Query Match 19.6%; Score 331; DB 4; Length 315;
 Best Local Similarity 28.6%; Pred. No. 1.5e-18;
 Matches 90; Conservative 65; Mismatches 140; Indels 20; Gaps 7;
 QY 13 VIGLFLVLIATAACGNNSSNS-SKESKDGVEIKHEGTTKPKPKRVVVLRYSFVDA 71
 DB 15 VAPLGLVALLSSCSTSSDSIQPEVASTGSVSEHANGTTEIPETPRVVVIDPSHLDA 74
 QY 72 LVALDVKPVGADNKNRIKPLRDKIGKTVSGTRKQPNLEISKLPDLIIADNNRH 131
 DB 75 LLALGITPVGATESGSENGFPAYLADELKUTSEVGLTSEPNLEKIALDPLIIGAKVRH 134
 QY 132 KGIYKDLNKIAPTTELKSPDGDYNNENIDAPKTIISKALGKEGKRLKEHDKTIEYKKE 191
 DB 135 EAIYDQSDIAPTVMSESGSGTNWNEQE-ITAAVKNKSEDMOKLISDLDRATELGE 191
 QY 192 ITMDKNQKVLPAVAAKSGLLAHPSNYYGQFLSQGFKEALSDDVTKGLSKYLKGPVLM 251
 DB 192 IGADQQTASVMVRFTDNFRLVGP-ETPSGVSLEQVGP----DLGERDNEY---SMWEL 242
 QY 252 NTEITLSQVNPMP--IWNKASNEPSLKELEKDPVWKLVNAVKNQVRIILDRDLWARS 309
 DB 243 SSFNGIGDGLIITYTPGSPATYTPKISEL-----WVDSPAVRQGRTYEFEDETWVG 297
 QY 310 RGLTSSBEMAKELVE 324

Db 298 IGVLGANEILDDLEE 312
 RESULT 35
 ABU25343
 ID ABU25343 standard; protein; 321 AA.
 XX
 AC ABU25343;
 XX 19-JUN-2003 (first entry)
 DT Protein encoded by Prokaryotic essential gene #10870.
 DE Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 KW Clostridium difficile.
 OS WO2002077183-A2.
 PN 03-OCT-2002.
 PD 21-MAR-2002; 2002WO-US009107.
 PF 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR N-PSDB; ACA29213.
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX Claim 25; SEQ ID NO 53267; 1766pp; English.
 PS The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC on a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained

CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 321 AA;

Query Match 19.3%; Score 327; DB 6; Length 321;
Best Local Similarity 30.9%; Pred. No. 3.2e-18;
Matches 99; Conservative 55; Mismatches 136; Indels 30; Gaps 12;

QY 11 LSVIGLLFVLIATAAGNNSSNSSKESKDGVEIKHEEGTTKVPKPKRVVVLVYSFVD 70
DB 11 LAILGLVTVF---ALGGKKNESKTSSENSNTIKITHLGETDVKLNPKVWVFPYSALD 67

QY 71 ALVALDVKP--VGIADNNKNNRIIPLRDKIGKITSVGTTRKOPNLEESIKLPDLIIADN 128
DB 68 TWDALGVAENLVGLPKASLPASLEK-YKDK--KYTDLGLGLKPEPDLEGIKSANPDIIII-N 123

QY 129 NRHKGIYKDLNKIAPTIELKSPGDYNE-----NIDAFKTSIKALKEBEGKRLBEHDKK 184
DB 124 GRQEDFYEQLSKIAPTISTSKDDKKYLESVKNNID---KIAKIFGVEEKANQOEFKIEKK 180

QY 185 IBEYKKEITMDKNQKVLPAVAASKGLLAHPSNSYVQGFLSQLGKKEALSDVTKGLSKYL 244
DB 181 IETLNKKVT-DKNLNTIMWNEGNLSVGEESRFSILYNSFGP-----ENKDKNIKESS 234

QY 245 KGPYLQMTETLSQVNPFRMTM-----TNKASSNEPSLKELEKDPVWKKLNNAVKNQVVD 299
DB 235 HQQ--NITPEYIAKQNPWFVIDRGIATGSDVKESSTAKSVLNNDIIKSMDAYKNDNII 292

QY 300 ILDRDLW-ARSRGLISSEEM 318
DB 293 YLDSPTWYVNDGGLTSLNKM 312

RESULT 36

ID ABU18359
ABU18359 standard; protein; 314 AA.

AC ABU18359;
DT 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #3886.

DE Antisense; prokaryotic essential gene; cell proliferation; drug design.

DE Antisense; prokaryotic essential gene; cell proliferation; drug design.

DE Bacillus anthracis.

DE WO200277183-A2.

DE 03-OCT-2002.

DE 21-MAR-2002; 2002WO-US009107.

DE 21-MAR-2001; 2001US-00815242.

DE 06-SEP-2001; 2001US-00948993.

DE 25-OCT-2001; 2001US-0342923P.

DE 08-FEB-2002; 2002US-00072851.

DE 06-MAR-2002; 2002US-0362699P.

DE (ELIT-) ELITRA PHARM INC.

DE Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

DE Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;

DE WPI; 2003-029926/02.

DE N-PSDB; ACA22229.

XX New antisense nucleic acids, useful for identifying proteins or screening

XX for homologous nucleic acids required for cellular proliferation to

XX isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 46283; 1766pp; English.

PS

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 314 AA;

Query Match 18.5%; Score 313; DB 6; Length 314;
Best Local Similarity 29.9%; Pred. No. 4.3e-17;

Matches 99; Conservative 65; Mismatches 135; Indels 32; Gaps 14;

QY 9 KILSVIGLLFVLIATAAGNNSSNSSKESKDGVEIKHEEGTTKVPKPKRVVVLVYSF 68

DB 4 KLFILFTMLVLSIVGSSQKESKAKEQPKTKV-VKHAKGEATIPVNPKEIVDLSGS- 61

QY 69 VDIALVDKVPVGIADNNKNNRIIPLRDKIGKITSVGTTRKOP--NLBEISKLPDLIIA 126

DB 62 TEELLLGLHGVGTGTA-NYKDKIQKHLTEKLDGVKAVGWYAPKVDLEAVTALKPDLIIL 120

QY 127 DNNRHGIYKDLNKIAPTIELKSPGDYNEIDAFKTSIKALKEEGKRLBEHDKKIE 186

DB 121 -NNRQLKIYDQLEKVAPTVLTNLEDR--GRKFEVKGKLFDEKKADKIADYDKKAD 176

QY 187 EYKKEITMDKNQKVLPAVAASKGLLAHPSNSVY-GQFLSQLGFKKALSDV--TKGLS 241

DB 177 SLSKIKKEKTKDDSPMFVAVT-----PQNFVYVYGSF-----GYGDIIFNDLKLPAKGTG 226

QY 242 KYLKGPIYQMTETLSQVNPFRMTMNTKASSNEPSLKELEKDPVWKKLNNAVKNQVDTL 301

DB 227 --LQQTWAVQSLEGLVAFQPDQMFIV-NFGGEADKVVYEDYKNKSAVWKNKAVKNNHVEV 283

QY 302 DRDLW-ARSRGLISS-----EEMAKELVELSK 327

DB 284 SNEVENTKAFNPIGKMDLIDIEIAKILAKNK 314

RESULT 37

ADC94889

ID ADC94889 standard; protein; 352 AA.

XX

AC ADC94889;

XX

DT 01-JAN-2004 (first entry)

XX

DE E. faecium protein sequence SEQ ID 4516.

XX Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;
 KW abdominal-pelvic infection.
 XX Enterococcus faecium.
 XX US6583275-B1.
 XX 24-JUN-2003.
 XX 30-JUN-1998; 98US-00107532.
 XX 02-JUL-1997; 97US-0051571P.
 XX 14-MAY-1998; 98US-0085598P.
 XX (GENO-) GENOME THERAPEUTICS CORP.
 XX Doucette-Stamm LA, Bush D;
 XX WPI; 2003-799836/75.
 XX N-PSDB; ADC91235.
 XX New isolated nucleic acid derived from *Enterococcus faecium* encoding an
 PT Enterococcus faecium polypeptide useful for detection, prevention and
 PT treatment of a pathological condition resulting from a bacterial
 PT infection.
 XX Example 1; SEQ ID NO 4516; 243pp; English.
 XX The invention relates to an isolated nucleic acid derived from
 XX Enterococcus faecium encoding an *Enterococcus faecium* polypeptide having
 CC one of 10 fully defined sequences given in the (or comprising 40
 CC sequential nucleotides chosen from any of the nucleic acids, its
 CC complement or sequences hybridising to it). Also included are a
 CC recombinant vector comprising the nucleic acid operably linked to
 CC a transcription regulatory element, a cell comprising the vector and a
 CC single-stranded probe comprising the nucleic acid. The nucleic acids are
 CC chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
 CC The nucleic acid is useful for diagnosing pathological conditions
 CC resulting from *E. faecium* bacterial infection (e.g. urinary tract
 CC infection, bacteraemia, endocarditis, wounds and abdominal-pelvic
 CC infection) and for screening drugs such as agonists and antagonists. The
 CC nucleic acid is useful for recombinant production of *Candida albicans* -
 CC derived peptides or antisense polypeptides. Pharmaceutical compositions
 CC and vaccines containing the nucleic acid are useful for preventing or
 CC treating *Enterococcus faecium* infections. The present sequence represents
 CC one if the disclosed *E. faecium* proteins.
 XX Sequence 352 AA;
 SQ
 Query Match 18.2%; Score 307.5; DB 7; Length 352;
 Best Local Similarity 31.3%; Pred. No. 1.4e-16;
 Matches 107; Conservative 65; Mismatches 123; Indels 47; Gaps 15;
 QY 5 VRGKILSVGLLVLIATAAG--NNSSNSSKESS---KDGVKHEEG-TTKVPKHP 58
 DB 40 VVGILLASFG-----AACGNNNTSQDTRKSTQTETTLITDSNGDQIEVPNPP 92
 QY 59 KRVVLEYSFVDALVALDV--KPGVIADNKKRIIFPLRDKIGKTYTSVGRKQPNLEBI 116
 DB 93 EKVVVFNGSLDTMDALGVGDKVVGATSS-----LPEYLSYKVESAGGIKEPDLEKI 147
 QY 117 SKLPDILLIADNNRHKGIVKDLNKIAPTEILKSPG-----DYNENIDAFKTSKALGKE 171
 DB 148 NQLQPDLLII--SGRQRPQKLSQTAFTIFL--SLDAKNPWSFQONVTA---LGEIFGKQ 202
 QY 172 BEGKKRLSEHKTEEVKKEITMDKNQKVPAAVAKSGLLAHPNSVYGQPLSQLGPKEA 231
 DB 203 BEATQLEELSSALDQTKKAAEA-TDKAALVTLVNEGQLSAYGSGSRFGFTHDLFGFEQA 261
 QY 232 LSDP-----VTKGLSKYLKPGYLOMNTETLSQVNPERFIM--TNKASSNEPSKLEKDPV 286
 DB 262 --DQIEASTHGQS-----VSVEYVLEKNPDILFVVDRTKAIGGDDSKDDISANEL 310

QY 287 WKKLNAVQNQRVDILDRDLWARSGGLISSEMAKVELVELSKK 328
 DB 311 VAQTNAGNQQIISLEPDVWYLSGGLESKMKIMIEDVNAQPK 352
 RESULT 38
 ABUL7461
 ID ABUL7461 standard; protein; 314 AA.
 XX AC ABUL7461;
 XX 19-JUN-2003 (first entry)
 XX Protein encoded by Prokaryotic essential gene #2988.
 XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX *Bacillus anthracis*.
 XX WO200277183-A2.
 XX 03-OCT-2002.
 XX 21-MAR-2002; 2002WO-US009107.
 XX 21-MAR-2001; 2001US-00815242.
 XX 06-SEP-2001; 2001US-00948993.
 XX 25-OCT-2001; 2001US-0342923P.
 XX 08-FEB-2002; 2002US-00072851.
 XX 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 XX N-PSDB; ACA21331.
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX Claim 25; SEQ ID NO 45385; 1766pp; English.
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of

KW antibacterial; gene therapy.
OS Staphylococcus epidermidis.
XX
PN US6380370-B1.
XX
PD 30-APR-2002.
XX
XX
PP 13-AUG-1998; 98US-00134001.
XX
XX 14-AUG-1997; 97US-0055779P.
PR 08-NOV-1997; 97US-0064964P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
PA
XX
XX
FI Doucette-Stamm LA, Bush D;
XX
XX WPI; 2002-381255/41.
DR N-PSDB; ABN93192.
XX
XX
PT Novel isolated nucleic acid encoding a Staphylococcus epidermis
FT polypeptide, useful for diagnosing and treating bacterial infections.
XX
XX Disclosure; SEQ ID NO 5492; 267bp; English.
XX
XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
CC antibacterial activity and can be used in gene therapy. The sequences can
CC also be used in the diagnosis and treatment of bacterial infections,
CC particularly S. epidermidis infections. The sequences can be used to
CC screen for compounds able to interfere with the S. epidermidis life cycle
CC or inhibit S. epidermidis infection. N.B. The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the USPTO web site
XX
SQ Sequence 351 AA;

Query Match 17.5%; Score 296.5; DB 5; Length 351;
Best Local Similarity 28.4%; Pred. No. 1.1e-15;
Matches 104; Conservative 70; Mismatches 131; Indels 61; Gaps 16;

QY 1 GVSVRGKILSLVIGLLFLVLIATAACGNNSSNSSKES-SKDGVYKHE----- 48
Db 1 GVAHMKTVLFLLSLVLTACSNSSNNNSISKKNSDSKETVTIKNSPASCKENNGS 60
QY 49 -----EGTTKVPKPRVWVLEYSFVDALVALDVKPVGIADDNK-----KNRIIKPLRD 97
Db 61 DKKKISNTVEVPKPKNAVLDYGDVLDL-----KELGVADKVKGLPKGENNQSLPKFLD 115
QY 98 --KIGKYSVQTRKQPNLEISKLKPDLI-IADNNRHKGIYKDLNKTAPIELKSPGDY 154
Db 116 EFKDDKYINTGNLKEVNFVKVSAKPDVIFISGRTANQKNLDEFKKAAPKAV-VYVGT 174
QY 155 NENI--DAFK--TISKALGKEEGKRLKLEHDKKIEYKKEITMDKNQKVLPAVAKSG 209
Db 175 DDNLIKDMKNNTENIGKIYDKEDKAKKINKDLDRKISDMKDK-TDFNKKVMYLLVNEGE 233
QY 210 LLAH-PSNSYVGQFLSQLGFKALSDDDVTKGLSKYLGKPYLQ-MNTETLSQVNP----- 261
Db 234 LSTFGPGRGGLVFDILGFKPA-----DDKVKSPHQNINNEYINKQNPVDILAM 285
QY 262 ERMFIMTNKASSNPSPSLKELEKDPVWKKLNKVNQRVDILDRDLWARSGLISSEMAKE 321
Db 286 DRGSVWGGKATTN-----QVLNKNVKNVAKNSHIYELDKLWYFSSG--SSTTTIKQ 338
QY 322 LVELSK 327
Db 339 IDELNE 344